

国际信息工程先进技术译丛

WILEY

# IPv6部署和管理

## IPv6 Deployment and Management


[美] Michael Dooley Timothy Rooney 著

董守玲 王昊翔 胡金龙 等译







机械工业出版社
CHINA MACHINE PRESS




国际信息工程先进技术译丛

# IPv6部署和管理

[美] Michael Dooley
Timothy Rooney 著
董守玲 王昊翔 胡金龙 等译



机 械 工 业 出 版 社

**图书在版编目（CIP）数据**

IPv6 部署和管理/（美）杜里（Dooley, M.），（美）鲁尼（Rooney, T.）著；董守玲等译. —北京：机械工业出版社，2015.1

（国际信息工程先进技术译丛）

书名原文：IPv6 deployment and management

ISBN 978-7-111-48725-8

Ⅰ.①I…　Ⅱ.①杜…②鲁…③董…　Ⅲ.①计算机网络－通信协议　Ⅳ.①TN915.04

中国版本图书馆 CIP 数据核字（2014）第 282690 号

机械工业出版社（北京市百万庄大街 22 号　邮政编码 100037）

策划编辑：刘星宁　　责任编辑：刘星宁

版式设计：霍永明　　责任校对：崔兴娜

责任印制：乔　宇

保定市中画美凯印刷有限公司印刷

2015 年 2 月第 1 版第 1 次印刷

169mm×239mm・12.5 印张・227 千字

0001—2500 册

标准书号：ISBN 978-7-111-48725-8

定价：58.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

| 电话服务 | 网络服务 |
| --- | --- |
| 服务咨询热线：010-88361066 | 机工官网：www.cmpbook.com |
| 读者购书热线：010-68326294 | 机工官博：weibo.com/cmp1952 |
| 010-88379203 | 教育服务网：www.cmpedu.com |
| **封面无防伪标均为盗版** | 金书网：www.golden-book.com |

本书首先阐述了为什么要发展和部署 IPv6，然后详细介绍了 IPv6 协议，以及 IPv4 与 IPv6 互通的技术。本书还详细描述了如何评估 IPv6 准备情况，如何进行 IPv6 地址规划，如何进行 IPv6 安全规划和管理规划，如何部署和管理 IPv6 网络，如何管理 IPv4/IPv6 网络等相关的策略和技术。最后，本书对 IPv6 和因特网进行了展望。

本书适合作为相关专业的工程技术人员学习下一代计算机网络协议 IPv6 概念和相关部署和管理技术的参考书，也可以作为计算机专业、网络工程专业和通信专业本科高年级的教材。

# 译者序

过去的二十多年里，IPv4 作为因特网的标准取得了辉煌的业绩。但随着物联网及大数据时代的到来，IPv4 因特网的缺陷已变得越加明显，如地址空间的不足、端到端通信质量得不到保障、网络配置管理困难等。相比于 IPv4，IPv6 具有如下优越性：良好的可扩展性、可靠的安全性、多样化的服务质量、易管理性、移动性；另外，还能很好地支持多播业务，提高网络的整体吞吐量。使用 IPv6 后，大量智能终端将有机会独立连接到因特网上，人们将生活在一个万物互联的世界中。近年来，越来越多的国家在 IPv6 研发中投入了巨大的精力和时间并进行部署和普及。我国在基于 IPv6 的下一代互联网方面开展了多项研究、试验和示范工程。2008 年 12 月 3 日，历时五年的中国下一代互联网（China Next Generation Internet，CNGI）项目取得阶段性的成果，建成并稳定运行全球第一个也是规模最大的纯 IPv6 互联网主干网。2011 年 12 月国务院常务会议指出我国已在基于 IPv6 的下一代互联网方面取得了阶段性进展，后续规模化商用和发展已具备良好基础，会议明确了今后一个时期我国发展下一代互联网的路线图和主要目标。

随着商用部署的持续推进，网络工程师、管理者和 IT 工作者迫切需要了解和学习 IPv6 基本概念并掌握 IPv6 部署与管理技术。本书《IPv6 部署和管理》正是为这样的需求而准备的。本书是一本系统、先进、实用的 IPv6 网络部署和管理的参考书，是从事的计算机网络技术研究、应用、开发和管理的网络工程师、管理者和 IT 工作者的良师益友。

本书包括 10 章和 1 个附录。第 1 章阐述了为什么要发展 IPv6，并从商业前景的角度阐明了各种组织团体应该部署 IPv6 的理由。第 2 章详细介绍了 IPv6 的协议特点及结构等基本概念。第 3 章详细介绍 IPv4/IPv6 共存技术，包括双栈、隧道、翻译技术等相关的技术。第 4 章详细介绍了如何评估现有网络的状况，升级或部署 IPv6 需要做什么准备。第 5 章详细介绍了制定 IPv6 地址规划的机制和技术。第 6 章从安全的视角探讨了 IPv4 和 IPv6 之间的差异，并突出了在更新安全策略时需要考虑的一些关键点。第 7 章介绍了网络管理的模型、范围、管理协议和管理功能。第 8 章介绍了 IPv6 部署管理过程。第 9 章介绍了管理 IPv4/IPv6 网络的策略和技术。第 10 章对 IPv6 和因特网进行了展望。附录 1 是 IPv6 准备情况评估表的模板。

董守玲、王昊翔、胡金龙组织并参加了本书的翻译和审校工作，参加翻

译的还有缪如倩、苏孟辉、张浩威、张铃启、陈伟健、滕菲、陈泽邦、刘荣波。

限于译者的水平，译文中难免有疏漏和错误，欢迎批评指正。

译　　者

2014 年于华南理工大学

# 原书前言

从具体阐述 IPv6 数据报结构的 RFC 2460 文章发表到现在已经有 14 年（2013）了。由 Setve Deering 和 Bob Hinden 撰写的这篇文章描绘了从 1990 年年初开始长达 8 年之久的关于“如何让 32 位的 IP 地址空间得到扩展”的辩论。在当时，对于 IPng（下一代 IP）有四个建议。我不打算将这些建议一一罗列出来，只想说明它们设定的功能差异巨大。曾经还有另外一条建议说接受 OSI 的无连接网络协议（Connectionless Networking Protocol，CLNP），这也引起了当时在因特网工程任务组（Internet Engineering Task Force，IETF）工作的那些一腔热血的工程师们的愤怒。这条建议也成为当时因特网工程任务组每次议程的首要议题。

在这些辩论之后，当时 IPng 工作组的联合主席，Deering 和 Hinden，于 1998 年记录下全部的辩论成果，然后将这些成果提交到互联网工程指导小组（Internet Engineering Steering Group，IESG），由 RFC 编辑来发布。当时我们许多人都期待着会立刻有人努力来实现这一协议。在网络快速发展的年代里，一直有着一个特殊的忧虑，那就是网络地址的使用率过快增长。新的互联网公司就像雨后春笋一样大量地冒了出来。在 IPng 辩论的同时，也进行着控制 IPv4 地址使用的努力，如重新诠释地址结构的每一位等的做法也在如火如荼地进行着。被称为无类域间路由（Classless Interdomain Routing，CIDR）的算法，通过允许使用任意的位数来区别网络号和主机号，让 IP 地址使用得更加充分和高效率。而且，自治系统（Autonomous System，As）的概念也被引入，它结合相关指标来阐明边界位置。另外，还改进了边界网关协议（Board Gateway Protocol，BGP），考虑用掩码来标识地址格式中网络号和主机号的扩展。再加上各地区的因特网登记处使用的异常严格的规则，IPv4 地址空间的消耗速度从根本上减缓了。IPv4 地址规划得如此好，以至于实现 IPv6 的压力慢慢地消散了。

网络地址转换（Network Address Translation，NAT）功能也被采用了。这个功能允许多部设备使用私有的网络地址，并且共同使用同一公有地址空间。在一个局域网中，NAT 技术通过使用端口号来映射公网地址与各个设备使用的私网地址。这个实践的成功吸引了电缆和无线通信设备的制造商，因为它们现在可以最大限度地让更多的设备来共享一个 IP 地址。NAT 技术大大增长了互联网服务提供商的因特网服务的注册用户数。

这些多样的措施延续了 IPv4 地址的使用，直到 2011 年 2 月互联网地址编码

分配机构（Internet Assigned Numbers Authority，IANA）在互联网名称和数字地址分配机构（Internet Corporation for Assigned Names and Numbers，ICANN）的赞助下召开会议，宣布该组织已经没有可分配的 IPv4 地址。各地区的因特网登记处（ARIN，LACNIC，RIPE-NCC，AFRINIC，APNIC）依然还有地址可供分配。但是不久，在 2011 年的 4 月 APNIX 的 IPv4 地址就耗尽了；2012 年 9 月，RIPE-NCC 也宣布它的 IPv4 地址已经耗尽了。一个以 IPv4 地址空间的交易市场已经形成，可是没能满足真正的需求。

物联网已经距离我们越来越近了。移动设备使用 LTE 技术传输数据，这需要端到端的通信能力。同样的，对于机顶盒，传感器设备，配备互联网功能的汽车，不计其数的家用和办公电器，还有那些可以嵌入在我们身体中的设备，也是这样的。解决这个问题唯一明智的方法就是在兼容 IPv4 地址的同时，实现 IPv6 地址。我们不能简单地“抛出一个开关”就能实现网上的所有设备从 IPv4 寻址到 IPv6 寻址的转换。这个过渡需要多年的时间。

这个长期的过渡使我们需要具备非常缜密的设计、细致的控制实现及周全的管理系统，从而能够同时处理当前网络和各种设备中的 IPv4 和 IPv6。我们不能为了简单化而把地址空间全部设计成单纯 IPv4 或者单纯 IPv6 的“孤岛”。可移动或便携式的设备会经常遇到 IPv4 和 IPv6 混合的环境。这对于那些缺少 IPv4 地址空间的因特网区域来说，确实是个不错的机会来使用纯的 IPv6 地址。复杂的运行环境不仅会包含经过 NAT 的 IPv4 地址，而且也含有 IPv6 的端到端传输。所以，网络工程师的书架（或者在笔记本、ipad、手机、云端、数码阅读器）上需要有一本由 Michale Dooley 和 Timothy Rooney 写的本书也就不足为奇了。

配置和网络管理是很难的，对于一个混合了两种 IP 数据报结构的环境，这类的操作更加困难。即便是一个普通的故障，如光纤被切断等，都会使两个协议产生错误信息。网络管理系统在过滤、关联、分类不同的错误信息和状态或者来自一个混合 IP 寻址环境的警告信息等功能上，都需要更加智能。IPv6 协议中较大数据报头部会导致分片，或者使为了避免分片阻塞而要发现最小数据报的过程复杂化。这里只有很少的问题需要回答。任何体系结构的系统只要考虑使用双协议栈的环境，都会发现本书是您的得力助手，是您所需建议的源泉。

现在开始这些实现是非常及时的。在这个十年的剩余几年中，我们会看到因特网会在许多方面产生巨大的变化和扩展，这些变化不仅表现在各个系统中相关设备的大量增长。

某些因特网服务提供商（Internet Service Provider，ISP）以“用户没有要求 IPv6”为借口而延缓了 IPv6 的实现。在我看来，用户不需要知道任何关于 IPv6 的知识。用户们都有一个合情合理的期望，认为他们的 ISP 会自动地实现双协议

栈，而不用用户自己去问。在 IPv4 地址全部耗尽之前，尽管使用了 NAT 技术，但不尽快实现双协议栈都是不负责任的做法。我们必须尽快地完成这个过渡期，实现全连接的 IPv6 网络。这并不意味着我们要抛弃 IPv4 的使用，只是我们需要 IPv6 网络能像现在 IPv4 做到的一样，而且我们现在就需要。

Vint Cerf

# 致　谢

我们要感谢 Vint Cerf 为这本书写了前言；我们感到很荣幸。我们也要感谢我们系列丛书的主编、在 IEEE 出版社工作的 Thomas Plevyak，当然还有 Michael Vincent 和 Jeff Schmidt，他们花了许多时间审阅本书，并且提供了许多非常有用的反馈和意见。

Michael：我也要谢谢我的家人，我的妻子 Suzanne，我的儿子小 Michale，我的女儿 Kelly，谢谢他们在我撰写这本书期间给予我深深的爱和坚定的支持，谢谢他们让我能专心致志地写书而不被分散注意力。同时我也不能忘记我的狗 Bailey，它总是来推我宠它一下，让我休息片刻。我要谢谢我的朋友和同事，我很高兴和世界上最好、最聪明的人一起工作，真心祝福他们。我要谢谢（不以特定的顺序）Karen Pell、Steve Thompson、Greg Rabil、John Ramkawsky、Alex Drescher、Brian Hart（又称 Billy Bond）、Bob Lieber、David Cross、Al Hilton。我要感谢原先由 Quadritek 领导的团队，早年的时候我有幸能和他们一起工作而定义和创造了 IP 地址管理的市场，特别是 Arun Kapur、Keith Larson、Leah Kelly。最后特别感谢 Joe D'Andere，他的领导能力对我的生活和事业都有着深远地影响。

Timothy：我也要谢谢我的家人，我的妻子 LeeAnn，我的两个女儿 Maeve 和 Tess，谢谢她们在我写这本书的期间给予我深深的爱和坚定的支持。我要感谢和我一起工作的同事们，同他们工作非常愉快，从他们的身上我学到了许多通信技术和 IPv6 的相关知识。他们分别是，Greg Rabil、John Ramkawsky、Andy D'Ambrosio、Alex Drecher、David Cross、Marco Mecarelli、Brian Hart、Frank Jennings。还要谢谢我在 BT Diamond IP、INS、Lucent 工作的时候的同事们。我在网络领域研究的形成期是在美国贝尔实验室，非常感谢 John Marciszewski、Anthony Longhitano、Sampath Ramaswami、Maryclaire Brescia、Krishna Murti、Gaston Arredondo、Robert Schoenweisner、Tom Walker、Ray Pennotti，特别感谢 Thomas Chu。

# 目　　录

# 第1章　IPv6实施的动力

## 1.1　因特网：一个成功的故事

因特网（Internet）已经产生很久了，从20世纪60年代美国国防部的一个有抗毁性的互相连接的网络，已经演化成一个全球性的通信方式。Tim Berners-Lee发明了万维网（WWW，Word Wide Web），定义了超文本格式，从而连接了不同的信息，并且这些信息可以方便地由浏览器获取。万维网加上那个简单的只需通过点击来进行的用户交互手段，让因特网从政府和科学家的实验室进入了寻常百姓家。Email是第二重要的因特网应用，它也促进了20世纪90年代中期对因特网服务的广泛应用。今天的因特网用户会发现这些丰富的信息和各式各样不同的应用程序在他们的生活中不可或缺。假如著名的网络应用，像Facebook、Youtube、Twitter、Google、Blogger，网购及新闻网站，甚至那些出色的电子邮箱服务突然消失了，大多数人会变得无所适从。

但是目前，因特网中丰富的信息和大量的应用程序也不是世界各地都能方便地访问的。图1-1所示的世界各地因特网渗透率，这是因特网世界统计（Internet World Stats）网站的统计数据，描绘了在2012年中期不同地区的因特网的渗透率，它是由当地因特网用户量占总人口比例来衡量的。只有1/3强的世界人口可以使用因特网。因特网渗透率在北美是最高的，超过78%；欧洲次之，约63%；在亚洲人口中，渗透率只有约28%。

再从另一个角度来看相同的数据，图1-2所示为世界各地因特网用户的比例。对比图1-1和图1-2所示，我们可以注意到，虽然在亚洲的因特网渗透率是28%，但是亚洲的因特网用户是最多的，有十亿之多，占世界因特网用户的45%。Internet World Stats网站估计世界因特网用户达到24亿。

因为世界因特网用户渗透率只有34%，看起来似乎还有很大的增长空间；而且，由于存在每个人需要多台不同设备的可能性，这个会产生对IP地址的大量需求。但是，什么样的环境会促进这样子的增长呢？最近一个世界银行的调查报告显示：在中低收入国家，每10%的互联网渗透率的增长对应1.12%的国家的平均经济增长（以国内生产总值衡量）[2]。另一方面，10%的宽带渗透率平均带来1.38%的GDP增长。报告还赞扬了宽带建设带来的社会经济福利，包括增加就业、增加创业机会、提供社会联系及提供公共基础信息服务。即使有些

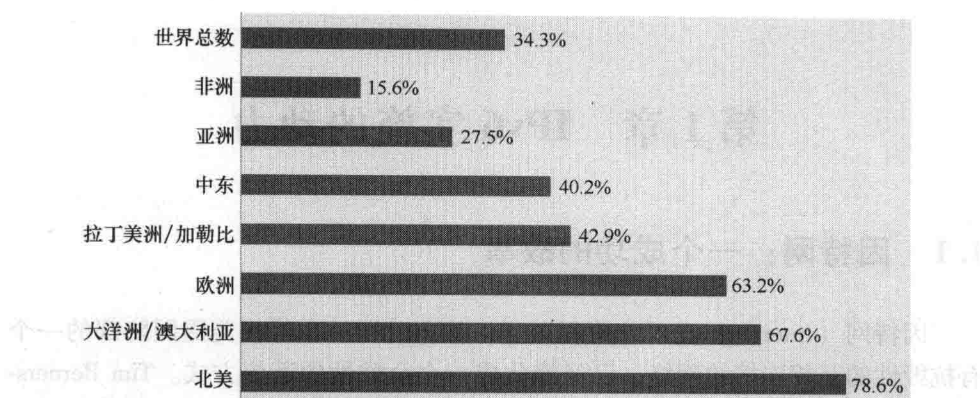



图 1-1 世界各地因特网渗透率[1]

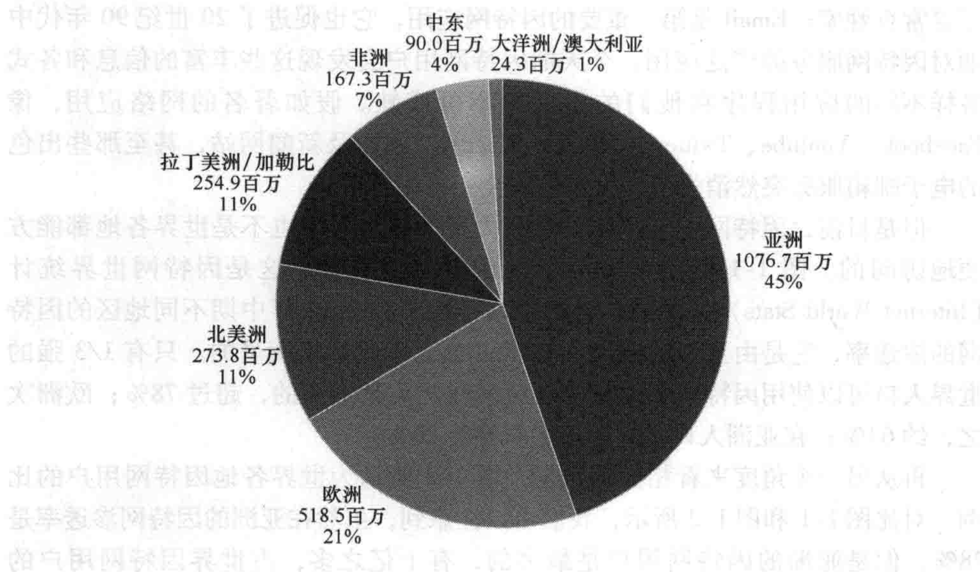



图 1-2 世界各地因特网用户的比例

国家会限制对于一些内容或应用的使用，经济和因特网的增长关系还是很难被忽视的。

在过去的十年显示强劲的增长趋势。图 1-3 给出了因特网用户和世界渗透率的增长情况。这看起来就是呈比例增长。表 1-1 给出了世界各地因特网用户增长情况，其中包括 2000 ~ 2011 年各地区的年复合增长率（Compound Annual Growth Rate，CAGR），世界平均复合年增长率是 18%。

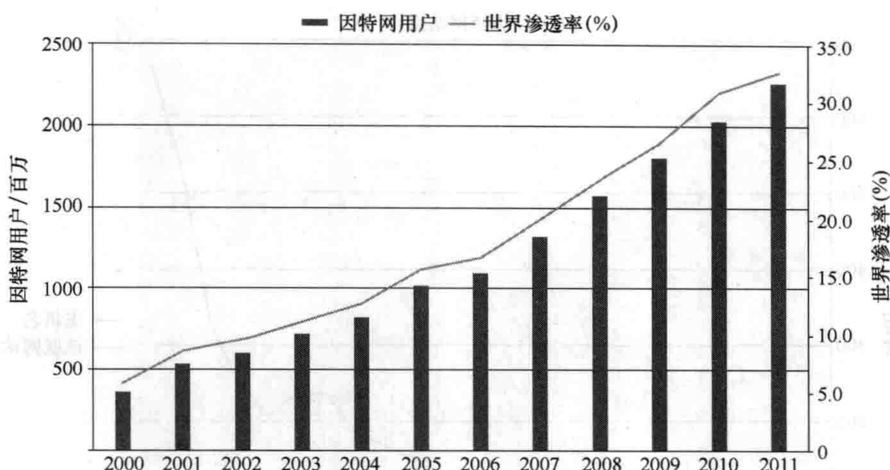



图 1-3　因特网用户和世界渗透率的增长情况

**表 1-1　世界各地因特网用户增长情况**

| 地　　区 | 2000/百万 | 2011/百万 | CAGR（%） |
|---|---|---|---|
| 非洲 | 4.5 | 139.9 | 37 |
| 亚洲 | 114.3 | 1016.8 | 22 |
| 欧洲 | 105.1 | 500.7 | 15 |
| 中东 | 3.3 | 77.0 | 33 |
| 北美洲 | 108.1 | 273.1 | 9 |
| 拉丁美洲/加勒比 | 18.1 | 235.8 | 26 |
| 大洋洲/澳大利亚 | 7.6 | 23.9 | 11 |
| 世界总数 | 361.0 | 2267.2 | 18 |

内容提供商或者“制造者（producers）”的数量在因特网上也在快速增长，这是通过新网站的数量来衡量的（根据英国 Netcraft 公司[3]（一家互联网研究与安全服务公司）提供的数据）。如图 1-4 所示，到 2012 年 6 月，发现的唯一网络主机名的总数达到近七亿，其中活跃网站（非模板（non-template），仅根据注册信息）将近有两亿。两种度量值都在过去两年中加速上涨。在某一个时间点以后也许新的想要拥有自己的网站和主机的组织机构只能注册 IPv6 地址来提供服务。

### 1.1.1　供应方面的问题

考虑到因特用户和内容提供商数量持续增长的历史、相对不高的渗透率和因特网增长带来的经济利益这系列因素以后，有理由相信因特网用户需求和生产者需求将会持续上升。不幸的是，如今 IPv4 地址分配能力不足以支持这种需求的增长。而 IPv4 地址枯竭之时，无论从哪点来看，唯一可以用来支持这种地

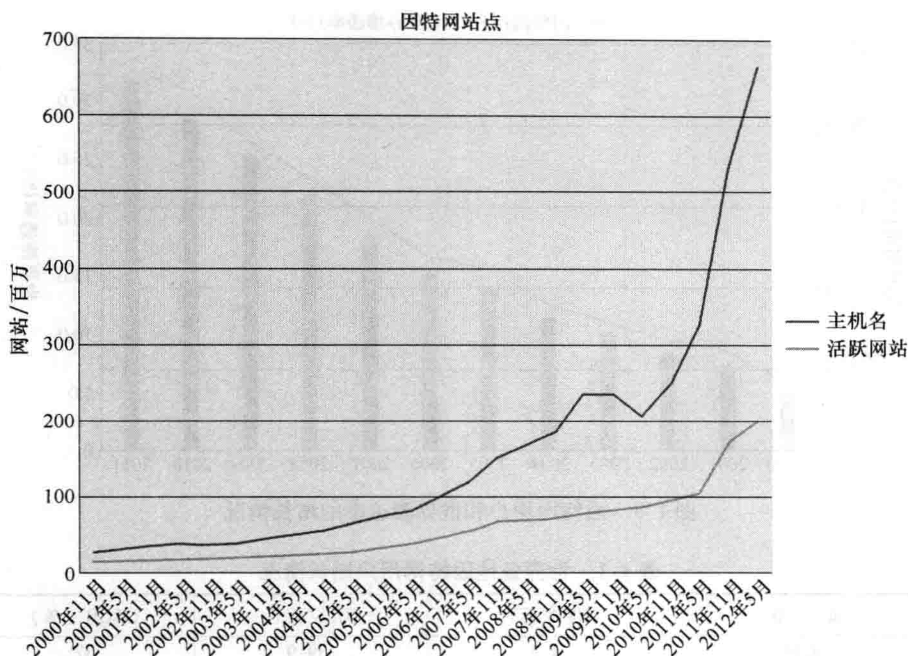



图 1-4 已测量的因特网网站数量[3]

址分配需求的协议只有 IPv6。

因特网名称和数字地址分配机构（Internet Assigned Names and Numbers Authority，IANA）在 2011 年 2 月 3 日宣布已经分配出最后一个地址空间给区域因特网注册管理机构（Regional Internet Registry，RIR），那一天其实就是所知的因特网末日的开始。图 1-5 所示的 IP 地址空间层次，解释了 IP 地址空间的“食物链（food chain）”。其中，IANA 把基础地址块分配给 RIR。IANA 是因特网名称和编号分配公司（Internet Corporation for Assigned Names and Numbers，ICANN）的一个部门，而 ICANN 本身是一个由世界各地成员组成的非营利性公益组织。IANA 是因特网域名的中心协调主体，管理着域名系统（Domain Name System，DNS）的根（root）域名和一些其他最高等级（top-level）的域名、因特网号码资源（IP 地址），以及协议任务（特定协议的参数，如动态主机配置协议（Dynamic Host Configuration Protocol，DHCP）的可选参数的任务）。

RIR 是一个负责将它们从 IANA 得到的地址空间在它们对应的地区进行地址分配的组织：

- AfriNIC（Africa Network Information Centre）——非洲地区。
- APNIC（Asia-Pacific Network Information Centre）——亚太地区。
- ARIN（American Registry for Internet Numbers）——北美地区，包括波多

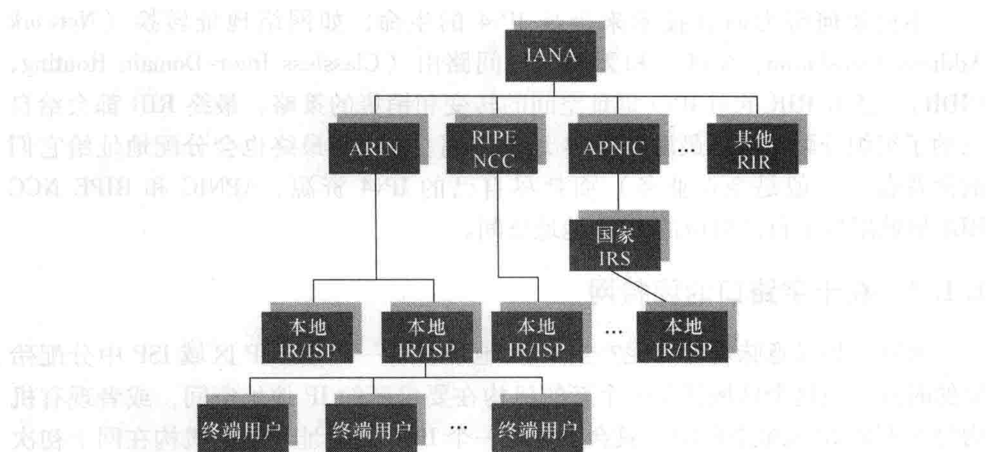



图 1-5　IP 地址空间层次[4]

黎各和一些加勒比海岛。

● LACNIC（Regional Latin-American and Caribbean IP Address Registry）——拉丁美洲和一些加勒比海岛。

● RIPE NCC（Reseaux IP Europeens Network Coordination Centre）——欧洲、中东和亚洲中部。

RIR 负责分配 IP 地址给国家因特网注册管理机构（National Internet Registriy，National IR）、当地因特网注册管理机构（Local Internet Registry，LIR）和互联网服务提供商（Internet Service Provider，ISP）。它分配 IP 地址时会遵循以下原则：

● 唯一性（Uniqueness），每一个 IP 地址必须是世界唯一的，以保证全球因特网选路。

● 聚合（Aggregation），分级的地址空间分配保证了对因特网上 IP 流量的合理路由。没有聚合，路由表会变得破碎不堪，这最终会造成因特网中的一个大瓶颈。

● 保护（Conservation），特别是 IPv4，但也包括 IPv6，地址必须根据实际的需求进行分配。

● 登记（Registration），一个公共可访问的 IP 地址分配记录，可以减少分歧并且可以在解决纷争时起到帮助。这个记录就叫做 WHOIS 数据库。今天，已经有很多 WHOIS 数据库。只不过现在这些数据库不再仅是由 RIR 运作，还可能由 LIR/ISP 运作。

● 公正（Fairness），主要根据地址的实际需求而非长期的计划来不偏不倚地分配地址。

不论如何努力通过技术来延长 IPv4 的生命，如网络地址转换（Network Address Translation，NAT）和无类别域间路由（Classless Inter-Domain Routing，CIDR），还有 RIR 的对 IPv4 地址空间的转变和销售的策略，最终 RIR 都会给自己的子组织分配出自己最后的 IPv4 地址空间。而 ISP 最终也会分配地址给它们的消费者（一般是企业业务）而耗尽自己的 IPv4 资源。APNIC 和 RIPE NCC RIR 早就用尽了自己对应的 IPv4 地址空间。

### 1.1.2 在十字路口的因特网

而这一切又意味着什么呢？当 IPv4 地址在某一给定 RIP 区域 ISP 中分配殆尽的时候，当这个区域任何一个新的机构在要求新的 IP 地址空间，或者现有机构要求追加 IP 地址空间时，就会被分配一个 IPv6 的地址。新的机构在网上初次出现时，它们将只能通过 IPv6 进行访问。而新的“纯 IPv6”组织持续加入网络，则会导致互联网组成由纯 IPv4 互联网慢慢转变成 IPv4 和 IPv6 的混合互联网。至于 IPv6 协议在互联网中的密度究竟会增长得多快，最后能增长到什么程度，还有待观察。

纯 IPv6 用户的增长应该最早会出现在亚洲。如图 1-6 所示，根据国际货币基金组织（International Monetary Fund，IMF）的统计结果，亚洲主要国家（特别是中国和印度）的经济在最近几年比世界其他地方增长得更迅速。从 2000 年到 2011 年，中国的年平均 GDP 增长率是 10.2%，印度的是 7.1%，而世界平均值是 2.7%。这样就相应增加了可支配的收入，以及政府在通信技术（如宽带和无线）上的基础设施投资。英国 Point Topic 有限公司分析认为，在 2012 年上半年世界上增加的宽带几乎一半在亚洲。英国 Point Topic 公司的报告和 Internet

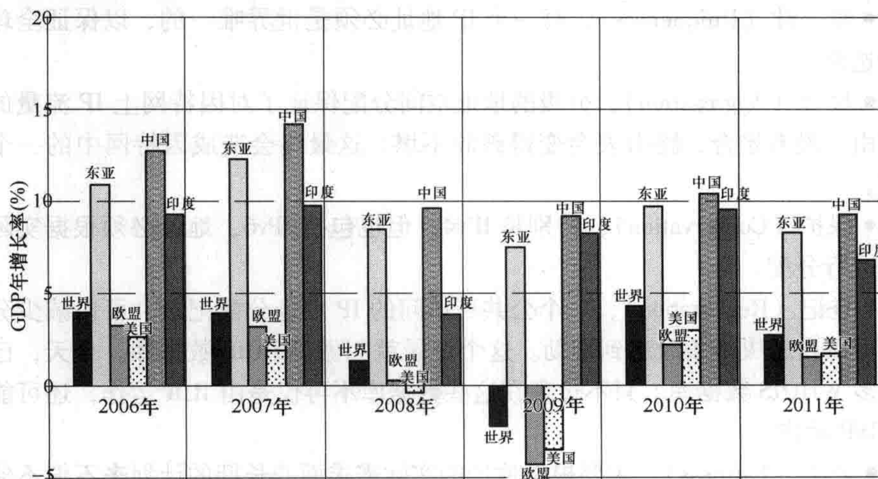



图 1-6 IMF 统计的 2006 年～2010 年世界及部分国家、地区的 GDP 增长情况

World Stats 网站的相呼应，Internet World Stats 网站的报告认为虽然亚洲有世界最多的互联网用户量，差不多有十亿，但是它的互联网渗透率是低于世界平均水平（33%）的，只有约 28%。

虽然在 2012 年增速有一些下降，但是因特网用户的增长率仍然十分高，这刺激了区域和世界的服务提供商实施转换网关、隧道代理等一系列类似技术。这些将在本书第 3 章详细描述（这些技术都是些保证 IPv4 和 IPv6 网络相互通信的技术——译者注）。根据所选择的方法，这些技术会为服务提供商提供用户驻地设备（Customer Premises Equipment，CPE），这个用户驻地设备会配置 IPv4 或 IPv6 地址并有能力访问任何 IPv4 或 IPv6 的目的地。服务提供商将必须使客户能够同时访问两种协议，以避免客户抱怨说自己是 IPv4 使用者而不能访问 IPv6 资源或说自己是 IPv6 使用者却不能访问 IPv4 资源。

### 1.1.3 使用哪一种因特网

现在可以很明确地知道，如今的因特网是在转型的关键点。在这个时候，也许你会问亚洲部署 IPv6 和自己有什么关系。其实理由很简单，无处不在的因特网都要面对这个问题。假设每一个有丰富 IPv4 地址的组织持续管理一个现存的纯的 IPv4 的网络，那么那些新增加的 IPv6 的用户将无法访问它们。与之相对，这些纯 IPv4 网络的用户也无法访问那些纯 IPv6 的网络内容。不幸的是，现在没有先天的 IPv4 与 IPv6 的转化机制，所以把一个 IPv6 的包发送给一个 IPv4 的网络服务器会导致这个包在到达服务器之前就被丢弃。因此一个分叉的互联网可能会衍生出两个不同版本的网络层。这将是最不幸的，更是目光短浅的。

不仅无所不在的因特网正处在关键点上，世界各国的竞争力和领导力也处在这样一个转变的关键点上。欧洲 RIPE NICC 组织定期的图表报告[8]显示了在自己广告中把 IPv6 加在前缀里的边际网关协议的自治区（BGP ASN）的比例。边际网关协议（Border Gateway Protocol，BGP）是在互联网主干网络上的路由协议。而每一个自治区编号（Autonomous System Number，ASN）代表一个组织。在 2012 年，所有地区都经历了超过 32% 的 IPv6 互联网路由数量的增长。到写本书时，世界上差不多有 15% 的组织（ASN）已经获取 IPv6 地址空间，并且部署了 IPv6，至少保证了其外部可达性。如果你的竞争对手在这之中，新兴的需要使用 IPv6 的市场，特别是亚洲市场，那么你究竟能多快地做出反应？总的来说，支持一个不断发展的现有的可通达世界各地的网络是 IPv6 部署时首要考虑的因素。

## 1.2 新兴的应用

除了无处不在的 IPv4/IPv6 网络之外，一些新兴的使用 IPv6 的应用将会给人

们的生活带来极大的变化。虽然技术上可支持 IPv4，这个“聪明的”应用有很高的移动性，很大的地址空间，并且可以根据当前网络位置自动分配地址，这些特征都不是 IPv4 能很好地支持的。虽然大多数为 IPv6 协议设计的特征与 IPv4 协议的有很多相接近的等价特征，使得可以重新应用在 IPv4 身上，但是地址扩展确实是 IPv6 的一个独一无二的优势。别的改进包括通过更有效的路由、地址自动分配、在网际的报文分段，来提升包路由效率，来提升移动性，通过简化 IPv6 头结构来提升路由性能。

从来没有人怀疑过 IPv6 全地址空间地址分配能力。而地址的自动配置和对移动性的支持也满足了主要由快速增长的移动设备驱动而持续增长的互联网的需求。但是，另一方面，这些能力也为一类新的应用提供了环境，如远程传感器可以监视、侦察和报告给中心应用处理。这种新兴应用是以机器到机器（Machine to Machine，M2M）的通信为基础的。M2M 定义了机器间通信的结构和方法，来收集大量的“大数据”信息，以供人类使用。

图 1-7 给出了基本的 M2M 的结构。从图的底部开始，最底部部署的是有特定功用的传感器（Sensor），它通过聚合器（Aggregator）把监控状态报告更新给 M2M 网关。一个为监控一个特定物体而使用的传感器群会包含一个 M2M 区域网关。聚合器可以被部署用于接收和处理来自 M2M 区域网关内部的传感器群的最新信息。在一些情况下，传感器会直接同 M2M 网关通信，而在其他情况下是由聚合器传递传感器信息给 M2M 网关，这个聚合器可能是一个传感器，也可能是别的如手机一样的设备。

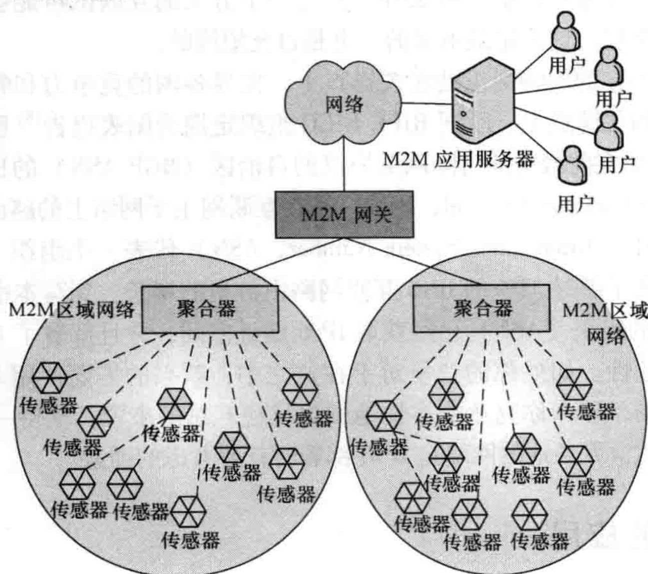



图 1-7 基本的 M2M 结构

M2M 网关会把聚合了的传感器的信息发送给 M2M 应用服务器，而这个服务器将处理的最新信息展示给应用使用者。M2M 服务器也可能对传感器报告的数据进行分析以发出警告，或者主动通知用户一些侦查到的事件。而这个“附加价值（valued-added）”处理能力，对于将来自千百万的传感器输入的数据预处理成一个可执行的按重要度划分清晰的监控元素视图是十分重要的。

随着尺寸、成本和能耗的不断减少，M2M 传感器设备可被用于各种不同的功能，并支持相应应用。M2M 传感器通过无线电与其他传感器、聚合器或者网关进行通信，提供传感器最新信息给相应应用进行分析。从 IP 层的角度来看，RFC 6568[9] 定义了“在低耗能无线个人空间网络上的 IPv6 的设计和应用空间（IPV6 over Low-power Wireless Personal Area Network，6LoWPAN）”。这个 6LoWPAN 结构和 M2M 对应得很好，LoWPAN 节点和传感器对应，当地控制节点和 M2M 聚合器对应。而网关接口则在 6LoWPAN 结构的 IP 层中贯穿一个 LoWPAN的边际路由器。

M2M 和 6LoWPAN 技术为服务提供商和传感器、应用开发者开辟了新的市场，扩展了对公用事业、政府当局、医疗机构和许多部门更多的服务。下面介绍一些 M2M 和 6LoWPAN 应用的例子。

- 智能应用（smart applications）。为许多智能资源管理系统和客户服务系统，提供一个对于一些没有察觉到的大量数据的中心视觉。这些系统包括：
  - 智能网格（smart grid）。根据需求自动调控电力、水、天然气等资源，减少资源浪费，节省消费者公用事业账单。
  - 智能汽车（smart cars）。在汽车里边的诊断的和使用的传感器，可以提供性能报告、发现并修复故障、提供用坏的组件的通知、推荐的服务检查及自动给出碰撞检测和报告。
  - 智能家庭（smart homes）。对房子远程监控，远程控制电、制热、制冷、光照、娱乐设施和出入。
- 城市和工业的监控和监视（Municipal and Industrial Surveillance and Monitoring）。物理访问的控制和监控、极端条件下的环境监控（如自然灾害、火灾、水灾）、结构监测和交通监测。
- 野外应用（Field Applications）。舰艇的管理、调度和舰载交通工具的远程信息处理。
- 卫生保健（Healthcare）。远程监视病人的重要生命的信号；诊断和药物管理；“身体区域网络”（body area networks）；严格的医用库存环境的监视，如血浆、器官的保存环境。
- 工业（Industrial）。工厂流水线监视、诊断程序、资源控制、供应链管理、作业管理和控制无线网络提供的可达性。

- 军事（Military）。战场专用网络通过不同士兵的传感器向军事指挥报告状态更新。

这些应用和别的与它们相似的应用通常需要部署成百上千甚至百万的传感器，这些传感器的度量和状态信息必须传送给中央应用服务器，以供处理和展示。M2M 结构是以一些十分常见的途径来支持这些功能的，这些途径包括网络接入（network access）、可靠通信（reliable communication）、安全保障（security）和集中管理（centralized management）。

可能会出现非常多的 M2M 设备，可以认为 IPv6 提供了这样一个逻辑网络层，它被“设计”（designed in）以提供非常大增长空间，特别是当新的 M2M 应用出现时可以避免之后对 IPv6 的升级。自动配置特性对 M2M 也十分有益，因为当传感器被部署并且被从省电模式唤醒的时候，它们可能会需要通过通告路由（router advertisement）确定自己连接的网络，并且通过接口 ID（Interface ID）获得 IPv6 地址。当成功通过冲突地址检测过程（duplicate address detection process）后，这个地址就是活跃（active）的了[⊖]。最后，选择 IPv6 协议在于与标准组相关的应用以及许多包括 IPv6 支持和 IPv4 支持的应用。

## 1.3 IPv6 商业案例

每一个组织必须要确定几个问题：是否需要部署 IPv6，什么时候部署 IPv6 和如何部署 IPv6。在“无作为”到“完全部署 IPv6”的连续区域中间有很多不同等级的实现方法。对于那些怀疑 IPv6 的实施是否必要的人，本书的建议是，至少要了解进行定量分析部署 IPv6 时需要做的事，这样在一些网络事件或者新闻刺激了你的领导，使他打电话给你让你尽快实施 IPv6 的时候，你能很快地实施。下面将会讨论这个高级过程（也就是分析 IPv6 实施需要做的事情的过程）。如果你打算开始向前进，对于之后工作的必要的深入讨论将在本书第 4 章展开。然而，最初的练习是有益于帮助你了解进行快速部署需要什么，以及大约需要多长时间的。

毫无疑问，因特网无论是在用户量上还是在内容量上都还在持续增长。这个增长必然会稀释几乎 100% 的 IPv4 网络密度，也必然会产生一个越来越混合的网络。同时支持 IPv4 和 IPv6 为你的组织保持网络的可达性并不是非常困难的，但也不是非常简单的。部署 IPv6 需要对现存的 IPv4 网络进行分析，确定 IPv6 部署范围，确定网络设备、应用或终端的升级或修改，以及管理整个项目

---

⊖ 地址状态更多的细节将在本书第 2 章讨论。

直到项目结束。本书的目的之一就是帮助你实现这整个过程。

在开始这个过程之前，需要你对所有的资源分配做出合理解释。从开始 IPv6 部署项目或该项目的发现和评估阶段，资本支出和费用的支付就必须要非常清楚。获取现有网络和计算系统的文档可以帮助你分析发现过程和评估过程甚至是整个部署过程的费用。

在以收益的提升、费用的降低和/或销售损失的减少来衡量带来的好处方面，在你的业务中还应该分析以下各点：

- 持续利润增长，特别是对于依靠 IP 连接性的服务提供商。
- 现存的世界性因特网，如果你的组织为世界各地的客户提供服务或产品。一开始不执行 IPv6 的机会成本是那些 IPv6 用户无法访问你的网站。由于因特网的发展很大程度上取决于 IPv6 的用户增长，你的组织将错失这些潜在的增长。此外你的内部用户也无法访问到 IPv6 网络上的资源。
- 如果你所在的组织属于信息技术产业领域，IPv6 的使用将产生更多的影响，可增加企业的竞争优势。
- 越来越多的雇员把他们自己的设备（Bring Your Own Device，BYOD）带来上班，很多现在和未来的便携设备都支持 IPv6，而且很多主流的操作系统都默认支持 IPv6。如果你的工作伙伴只有 IPv6 的地址空间，那么你至少应该想办法为和这种地址空间进行联系而支持 IPv6。
- 在拥有用户终端设备 IPv6 的支持后，其 IPv6 通信量的网络可见性。对于 IPv6 原生和隧道网络及来自外部使用 IPv6 的探查和攻击的察觉和可视化，对于网络安全是十分必要的。
- 使用 IPv6 特有的特征来支持新型应用，特别是移动性和自动配置。
- 为 IT 或者运作团队提供一个有趣而又有挑战的工作环境。据我所知，管理一个 IPv4/IPv6 环境肯定比管理一个单独协议的网站要复杂，但这也对雇员的知识增长和事业发展有益。
- 由于规则和法律要求而支持 IPv6。

你也许希望根据 IPv6 网络的密度来描述机会成本。例如，某个时间 IPv6 网络用户和网址在互联网上达到 20%，这也许可以代表一个足够大的人口数量，来证明部署 IPv6 并且与包括那 20% 的整个因特网相通的正确性。这仅是一个你们组织需要做的决策罢了。但无论这个密度是 1% 还是 99%，有一个在一个特定时间开始部署 IPv6 的计划对你来说都是十分重要的。

基本的项目授权过程如图 1-8 所示，它是从定义这个项目的目标开始的。一个或者多个上面提到的好处可以作为你这个部署过程的目标，它也可以帮助你关注部署实施的范围。有了你的目标、范围、对网络文档的高等级的评估及这本书概述的过程，你应该可以预测出资源费用和时间期限。根据你的组织的

不同，对项目具体批准的程度的期望也会不同。也许你只是希望批准发现和评估阶段之后重新遍历整个项目，那之后再考虑整体的批准。

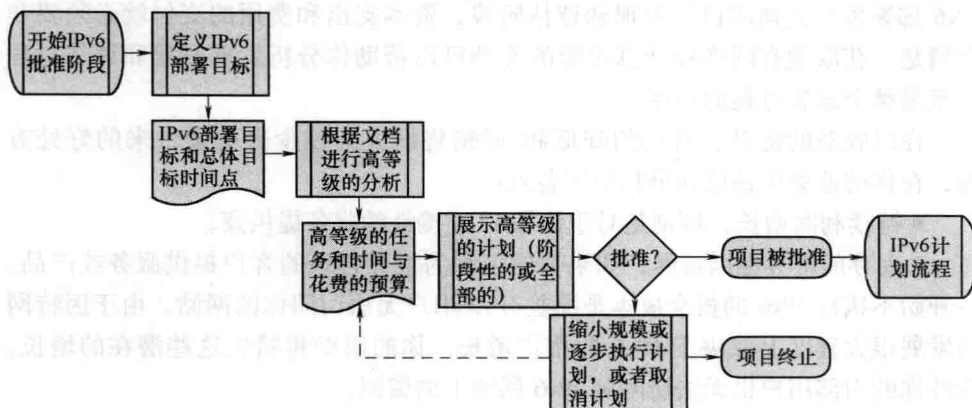



图 1-8 基本的项目授权过程

根据你决策的准则，评估你的网络、找出差距和创建项目计划，无论是对立刻执行部署或者是对延迟执行部署都有帮助。本书描述的整个部署过程包括五个基本步骤，如图 1-9 所示。本章已叙述了批准过程，这是一个需要对部署目标、部署范围、部署计划、部署预算和 IPv6 部署裨益的基本定义。本书下面很多内容将描述下一个阶段——计划阶段。这个阶段对于减少意外发生，并让整个部署过程更加顺畅十分重要。本书对计划阶段的四个核心方面各分一章描述，内容包括：网络和计算基础设施的评估和计划、IP 地址的计划、安全计划，以及网络管理计划。一个有效的计划之后是部署实施阶段，这个阶段包括初始测试和在部署运作后的验证过程。

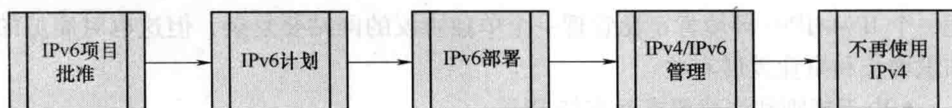



图 1-9 整个部署过程

一旦部署完成，除了一些小的修改和小的添加，对 IPv4/IPv6 网络的管理过程和单独管理 IPv4 网络的一样。在某一个时间点，也许是可预见的未来，IPv4 将会退休。现在这个时间点很难预计，但是某一天这一定会发生，虽然不大可能会在今后的 20 年内发生。

在开始对计划过程各个核心方面详细描述之前，十分有必要先了解 IPv6 和那些在 IPv4 网络里实施 IPv6 所用到的技术和策略。本章第 2 章和第 3 章先讨论这两方面内容。之后的章节将会描述计划过程、配置过程和管理过程。

# 第2章 IPv6 概述

IPv6[㊀]由 IPv4 演变而成，但是 IPv6 却并不与 IPv4 兼容[㊁]。这种不兼容使得 IPv6 需要在 IPv4 上另外进行部署。本书第3章描述了多种使得 IPv4/IPv6 兼容的技术。IPv6 建立的一个主要目标就是要根据过去 20 年使用 IPv4 的经验对 IPv4 进行重新设计。IPv4 对现实世界中的应用支持从一开始就被设计进了 IPv6，这包括对安全性、多播、移动性和自动配置的支持。

IPv6 与 IPv4 最显著的不同点是 IPv6 对 IP 地址字段进行了极大的扩充。这是因为 IPv4 使用的是 32 位地址，而 IPv6 使用了 128 位。一个 32 位地址空间可以提供最多 2 的 32 次方，约 42 亿个地址；而 128 位地址空间可以提供最多 2 的 128 次方，约 $3.4\times10^{38}$个地址[㊂]。下面用一些实际情况了解一下这个庞大的数字吧，例如：

- 对于拥有 65 亿人口的地球来说，平均每人可以拥有 $5\times10^{28}$个 IP 地址。
- 对于地球表面来说，平均每英寸土地都可以分配 $4.3\times10^{20}$个 IP 地址。
- 离我们最近的仙女座星系，大约有 250 万光年的距离，如此长的距离，每纳米上可以分配到 1400 万个 IP 地址。

如同 IPv4 一样，由于低效的子网划分，并不是每一个 IPv6 地址都被充分利用，但是一点点地址的浪费对于这么庞大的 IPv6 地址容量来说可以说是微不足道的。除了这看起来用之不竭的 IP 地址，IPv6 和 IPv4 也有许多相似的地方。例如，在最基本的层次，“IP 数据报”的概念也同样适用于 IPv6；又如 IPv4 中数据报头、数据报内容、协议层、数据报路由、CIDR 分配、互联网控制报文协议（Internet Control Message Protocol，ICMP），多播寻址等，也是如此。

## 2.1 IPv6 主要特性

IETF 尝试将 IPv4 演化成 IPv6。这种从 IPv4 迁移到 IPv6 的演化策略计划使

---

㊀ IP 版本 5 从来没有作为 IP 的官方版本实现，IP 报头的版本号“5”分配给表示携带一个称为 ST 的实验的实时流协议的包，即因特网流协议，如果你想了解更多有关 ST 的内容，请查看本书参考文献 [10] RFC 1819。

㊁ 本章部分信息基于本书参考文献 [11]。

㊂ 我们用美国的 undecillion，即 $10^{36}$，而不是英国的定义 $10^{66}$。

得 IPv6 提供许多新特性，这些新特性建立在使 IPv4 如此成功的基础概念上。IPv6 的主要特性包括以下几项：

- 扩展寻址——128 位分层分配的地址范围（如本地链路和全球）去提升扩展性。
- 路由——强大的分层路由，支持路由聚合。
- 性能——简单（不可靠、无连接）数据报服务。
- 可扩展性——新的灵活的扩展头为新头类型提供更好的扩展性和更高效的路由。
- 多媒体——流标签头字段促进服务质量（Quality of Service，QoS）支持。
- 多播——强制代替广播。
- 安全——内置的身份验证和加密，尽管不是强制的。
- 自动配置—— IP 设备的无状态及有状态地址的自动配置。
- 移动性——移动 IPv6 支持改进的网络路由和漫游。

## 2.2 IPv6 报头

IPv6 报头如图 2-1 所示。可以看到源地址和目标地址长度是 IPv4 的四倍，而 IPv6 报头的大小仅是 IPv4 的两倍。IPv6 报头字段如下：

版本（Version）。互联网协议，这里版本为 6。

流量类别（Traffic Class）。这个字段类似 IPv4 的服务类型/分区服务（Differentiated services）字段，用来区分流量类型或者流量的优先权，以此来请求响应的路由服务。

流标签（Flow label）。可以根据源地址和目标地址从“流”中区分报文的归属。这是为了在一个通信对话中保证高效和一致的路由服务。

有效负载长度（Payload length）。这表示在基本 IPv6 报头之后的部分还有多少个字节，显示了 IPv6 的负载长度。如果有扩展报头（Extension Head）的话，也是载荷的一部分，并且计算在负载长度之内。

下一个报头（Next Header）。这个字段指明了当前报头之后还有哪种扩展报头。这可能是一个上层协议的报头（如 TCP、ICMPv6），或者一个扩展报头。扩展报头可以用来提供一些额外的信息，如路由、分段、选项和一些有需要时与报文关联的其他参数，并非像 IPv4 一样要求全部报头都包含这些字段，而且参数的编码就跟那些已有的 IPv4 头字段中的编码一样。

跳数限制（Hot Limit）。跟 IPv4 中的生存时间（Time To Live，TTL）字段一样，这个字段指出了这个报文在被丢弃之前拥有多少跳数，每一次路由转发该报文时都会对这个字段进行递减操作。

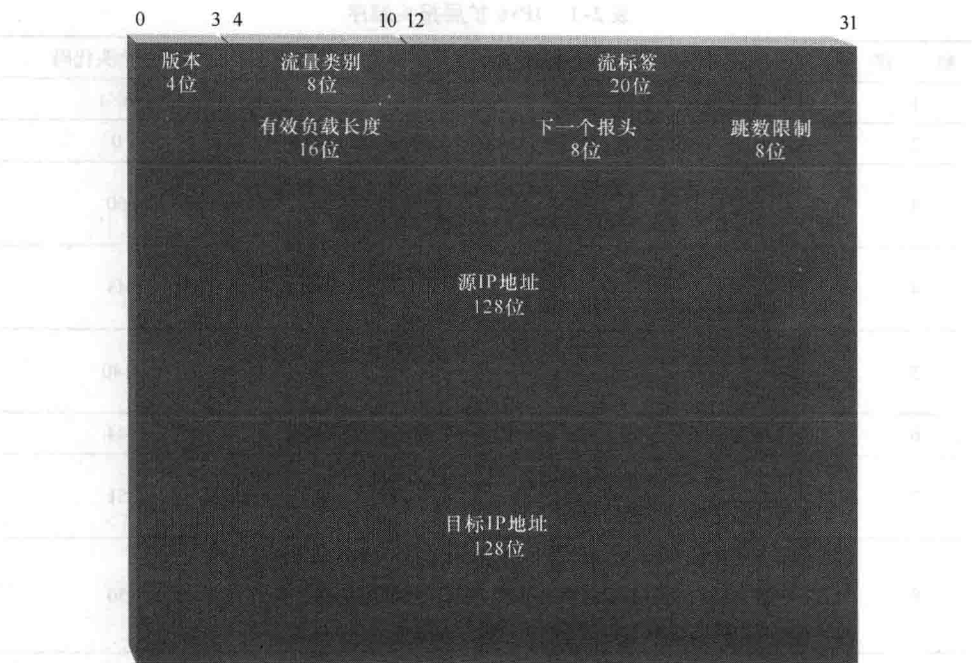



图 2-1　IPv6 报头

源 IP 地址（Source IP Address）。包发送者的 IPv6 地址。

目标 IP 地址（Destination IP Address）。报文预期接收者的 IPv6 地址。

### 2.2.1　IPv6 扩展报头

下一个报头（The Next Header）字段是 IPv6 提供的一种方式，以尽量减少 IPv6 报头的开销，同时还能够用于串联其他因某些目的被需要的报头，如图 2-2 所示。图中，左侧的主要 IPv6 报头结构表明下一个报头（Next Header）的值为 0，之后是逐跳报头（Hop-by-Hop header），接着是验证报头（Authentication Header），再然后是 TCP 报头，最后是传送给上层的负载数据。




图 2-2　IPv6 扩展报头

不同的报头类型具有不同的优先级顺序，以最小化 IP 包路由路径上的路由器在深入分析 IPv6 报头时对资源的要求。在 IPv6 报文经过的路径上的每一个节点，包括路由器，需要检查报文是否存在逐跳选项、路由报头和 shim6 报头。因此所需的 IPv6 扩展报头顺序见表 2-1。

**表 2-1 IPv6 扩展报头顺序**

| 顺　　序 | 报头类型 | 下一个头代码 |
|---|---|---|
| 1 | 基本 IPv6 报头（上面描述） | *N/A* |
| 2 | 逐跳选项——必须由报文所经过的路径上的每个节点检查 | 0 |
| 3 | 目标地址选项——参数用于处理出现在基本报头目标地址字段第一个目标地址及其随后路由报头中列出的目标地址 | 60 |
| 4 | 路由报头——最初定义的源路由（0 型路由报头）已经被弃用，1 型并没有使用，2 型则用于移动 | 43 |
| 5 | *SHIM6* 多宿主报头——用于在多宿主环境提供连通性（连接到多个 *ISP*） | 140 |
| 6 | 分段报头——关于被分割成多个部分而发送的报文的信息 | 44 |
| 7 | 验证报头——提供完整性校验值（Integrity Check Value，*ICV*）用于验证报文的完整性和源头 | 51 |
| 8 | 封装安全负荷（Encapsulating Security Payload，*ESP*）——提供了一个混合的安全服务包括数据来源认证、数据完整性、具有序列码完整性的防重放和有限的机密性（加密） | 50 |
| 9 | 移动报头——移动 *IPv* 6 信息 | 135 |
| 10 | 目标地址选项——用于处理最终目标地址的参数 | 60 |
| 上层 | *TCP*——传输控制协议 | 6 |
|  | *UDP*——用户数据报协议 | 17 |
|  | *ICMPv* 6——*IPv* 6 因特网控制报文协议 | 58 |

新的扩展报头可能用 IETF RFC 过程规范化，尽管 IETF 推荐在已有的逐跳报头增加新的逐跳扩展选项和在已存在的目的地址报头中添加目的地址选项。这会降低报文丢失的风险，因为由于安全策略无法识别新定义的下一个报头值。

## 2.3 IPv6 寻址

IPv6 地址的类型有 3 种。跟 IPv4 一样，这些地址识别为接口而不是节点。所以，可以通过任一接口寻址一个拥有两个接口的打印机。打印机可以根据任意一个接口到达，但是打印机本身没有一个回环地址⊖。当然，对于最终用户尝试访问一个节点，DNS 可以隐藏这个细节使得一个主机名可以映射到多个接口

⊖ 许多路由器和服务器产品通过软件回环地址支持“盒子地址（box address）”的概念，这个回环地址不要和 127.0.0.1 或 :: 1 回环地址混淆。这使得报文可以到达设备的任一接口。

地址。

**单播**，单个接口的 IP 地址。这类似一个 IPv4 主机地址的常见的解释（非多播/非广播/32 位 IPv4 地址）。

**任播**，一组接口的 IP 地址，通常属于不同的节点，其中的任何一个都可以是接收者。发往任播地址的报文会被路由到配置有任播地址节点的最近的接口（根据路由表度量值）。它的意思是发送者并不关心报文会被哪一个特定的主机或者接口接收，但是分享共同任播地址的其中一个会接收这个报文。任播地址与单播地址一样被分配相同的地址空间。因此，不能直接区分出单播地址和任播地址。任播地址常用来提供最近的路由到目标服务，如通过使用一个共享 IP 地址的 DNS。这提供了简化客户端配置的好处，使得可以总是使用相同的任播 IP 地址来查询一个 DNS（不管你的客户端连接的网络在哪里）。

**多播**，一组接口的 IP 地址，通常属于不同的节点，其中任意一个都可以作为接收者。这一点很像 IPv4 中的多播。与 IPv4 不同的是，IPv6 不支持广播。作为替代，应用程序，如 DHCP，在 IPv4 中使用广播而在 IPv6 中使用多播。

一个设备的接口可能有多个不同类型的 IP 地址。IPv6 也定义了一个本地链接地址范围来唯一标识接口使其附属于一个特定的链接，如 LAN。新建的范围域可以被用来管理每个站点或者每个组织，稍后将会在本章讨论。

## 2.3.1 地址符号

回想一下，IPv4 地址是在点分十进制格式下，分成 4 个 8 位段，表示 32 位地址，每一个都被转换成十进制，然后用“点”分隔。如果觉得记住一连 4 个数字都困难，那么 IPv6 会更加难。IPv6 地址不以点分十进制进行表达，它们使用冒号分隔的十六进制格式。从位的角度来说，128 位的 IPv6 地址分为 8 个 16 位段，每个段转换成十六进制，然后用冒号分开。每个十六进制的“数字”代表每一个十六进制数（0 ~ f）到 4 位二进制映射。每一个十六进制数对应于二进制四位的值为

| | | | |
|---|---|---|---|
| 0 = 0000 | 4 = 0100 | 8 = 1000 | c = 1100 |
| 1 = 0001 | 5 = 0101 | 9 = 1001 | d = 1101 |
| 2 = 0010 | 6 = 0110 | a = 1010 | e = 1110 |
| 3 = 0011 | 7 = 0111 | b = 1011 | f = 1111 |

将 128 位的 IPv6 地址从二进制转换十六进制后，IPv6 地址每个 16 位的段换算成 4 位的十六进制数并用冒号分隔开。这里用术语 nibble 代表 4 位二进制数或者 1 位十六进制数，用术语 sedectet 表示 16 位二进制数或者 4 位十六进制数。因此，有 8 个以冒号分开的 sedectet 值，呈现出的 IPv6 地址如图 2-3 所示。

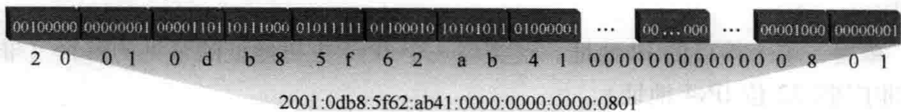



图 2-3 IPv6 地址（二进制转换为十六进制）

并不像处理以点分隔、0～255 的 4 个十进制组的 IPv4 地址一样，IPv6 地址由以冒号分开的 8 个在 0000～ffff 的 sedectet 值组成。有下面两种缩写形式可以表示 IPv6 地址：

第一种，在冒号之间的一个 sedectet 中出现前导零时，这个sedectet可以被简化。因此，图 2-3 的地址可以被简写为

2001：db8：5f62：ab41：0：0：0：801

第二种，是一种缩写形式，使用双冒号代表一个或者多个零 sedectet 连续集。使用这个形式的缩写，上面的地址可以进一步缩写为

2001：db8：5f62：ab41：：801

请注意，一个地址只能用一次双冒号表示。因为总共有 8 个 sedectet 地址段，可以很容易地计算出一个双冒号表示多少个 0。但是，如果有多个双冒号就可能会得出不同的答案。

考虑地址 2001：db8：0：56fa：0：0：0：b5。可以用下面任意一种形式的缩写地址：

2001：db8：：56fa：0：0：0：b5 或 2001：db8：0：56fa：：b5

在第一种情况下，可以很容易计算出双冒号表示一个 sedectet（8 减去 7）。在第二种情况下也可以很容易得出双冒号表示 3 个 sedectet（8 减去 5）。如果将此地址缩写成 2001：db8：：56fa：：b5，就不能明确解码，它可能代表下面任一地址：

2001：db8：0：56fa：0：0：0：b5
2001：db8：0：0：56fa：0：0：b5
2001：db8：0：0：0：56fa：0：b5

因此，要求一个 IPv6 地址中只能出现一次双冒号。根据 RFC 5652《推荐的一种 IPv6 地址文本表示》（本书参考文献［14］），正确的缩写是缩写最长的连续 0 段或有相同段数的 0 sedectet 时缩减第一段，示例 2001：db8：0：56fa：：b5 为正确的缩写形式。顺便说一句，RFC 5652 还规定使用小写字母的十六进制数。

你可能会注意到 IPv6 地址后缀为百分号后面跟一些数字或者文本，如 fe80：：9848：：e2f1：6d42 a87 % 11 或 a87 fe80：：9848：：e2f1：6d42% eth0。这个百分比符号划出了 IPv6 地址范围有时候被称为“scope ID”或“zone”。scope ID 值是由本地主机的操作系统通过识别地址的网络拓扑范围来定义的。例如，按照本地接口、本地连接、全球地址或者管理要求来定义拓扑范围。例如，

前文描述的 IPv6 地址范围可以定义一个本地链路地址范围的本地接口。拓扑区域（zone）的格式与组合是由设备的操作系统来定义的。

## 2.3.2 地址结构

IPv6 地址分为以下三个部分：

（1）全局路由前缀是类似一个 IPv4 网络号，并用于通过路由器的数据报转发到具有对应的网络前缀的路由器。例如，一个 ISP 的客户可能会被分配一个/48 大小的全局路由前缀，所有通往这个客户的数据报将会包含相应的全局路由前缀值。在这种情况下，如图 2-4 所示，$n=48$。当表示一个网络，全局路由前缀在斜线后被写入，紧接着的是网络的大小，称为前缀的长度。假设 IPv6 地址为 2001：db8：5f62：ab41：：801，存在/48 全局路由前缀，这个前缀地址将表示为 2001：db8：5f62：：/48。

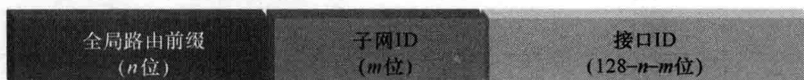



图 2-4 IPv6 地址结构[15]

请注意，在 IPv6 中使用的“网络斜线前缀长度”CIDR 表示法与 IPv4 使用的类似。与 IPv4 一样，前缀长度以外的位表示为零值的网络地址（本例中的第 49～128 位），用双冒号终止。

（2）子网 ID 提供了一种方法来表示组织内的特定子网。具有/48 的 ISP 客户会使用 16 位二进制来表示子网 ID，它提供了 $2^{16}$、共计 65535 个子网。在这种情况下，如图 2-4 所示，$m=16$。这就剩下 $128-48-16=64$ 位给接口 ID 了。

（3）接口 ID 表示报源或收件人的接口地址。正如之后将会讨论的，全局单播地址到目前已经分配使用的空间需要一个 64 位的接口 ID 字段⊖。

这个 IPv6 地址结构将一个网络 ID（由全局路由前缀、子网 ID 组成）从接口 ID 中分离出来。该结构的特性可以使一个设备保留相同的接口 ID 同时，独立地连接到网络，有效地根据接口 ID 区分“你是谁”，根据网络前缀分辨“你在哪”。正如所看到的，本公约方便了地址的自动配置，虽然不是没有隐私问题。但是这方面的讨论现在超前了一点，所以后退一点儿看一看目前 IPv6 地址空间分配情况。这是由 IANA 来负责管理的。

## 2.3.3 IPv6 地址分配

目前被 IANA 分配的 IPv6 地址空间见表 2-2，下面将进一步讨论。这些分配

⊖ 除了点到点路由器，各链路可以在每端用 1 位或者用一个 127 位的网络前缀[16]。

的 IPv6 地址占不到 14% 的地址空间。

表 2-2 目前被 IANA 分配的 IPv6 地址空间[17]

| IPv6 前缀 | 二进制形式 | IPv6 空间的相对大小 | 分配 |
|---|---|---|---|
| 0000::/3 | 000 | 1/8 | IETF 保留——未指定地址（::）和回路地址（::1）在这块分配 |
| 2000::/3 | 001 | 1/8 | 全局单播地址空间 |
| 4000::/3 | 010 | 1/8 | IETF 保留 |
| 6000::/3 | 011 | 1/8 | IETF 保留 |
| 8000::/3 | 100 | 1/8 | IETF 保留 |
| a000::/3 | 101 | 1/8 | IETF 保留 |
| c000::/3 | 110 | 1/8 | IETF 保留 |
| e000::/4 | 1110 | 1/16 | IETF 保留 |
| f000::/5 | 1111 0 | 1/32 | IETF 保留 |
| f800::/6 | 1111 10 | 1/64 | IETF 保留 |
| fc00::/7 | 1111 110 | 1/128 | 唯一的本地单播地址 |
| fe00::/9 | 1111 1110 0 | 1/512 | IETF 保留 |
| fe80::/10 | 1111 1110 01 | 1/1024 | 本地链路单播地址 |
| fec0::/10 | 1111 1110 11 | 1/1024 | IETF 保留 |
| ff00::/8 | 1111 1111 | 1/256 | 多播地址 |

#### 2.3.3.1 ::/3——预留空间

前缀 $[000]_2$ 的地址空间是现在 IETF 预留的地址空间。在这个空间的地址拥有独特意义的包括未指定（::）地址和回路（::1）地址。本书参考文献 [15] RFC 4291《IPv6 寻址体系结构规范》，要求所有单播 IPv6 地址，除了那些在预留地址空间内的地址（即除了那些以::/3（$[000]_2$）开始的）必须使用 64 位接口 ID 字段。这个接口 ID 字段必须利用修正的 EUI-64⊖算法映射接口的第二层或硬件地址到一个接口 ID。因此，在地址::/3 内的地址空间可以具有任意长度的接口 ID 字段，跟 IPv6 单播地址空间其他那些必须用 64 位接口 ID 字段的部分不一样。

#### 2.3.3.2 2000::/3——全局单播地址空间

到目前为止，全局单播地址空间分配了 2000::/3，代表 $2^{125}$（大约 $4.25\times10^{37}$）个 IP 地址。鉴于在上述 IPv6 寻址结构（RFC 4291）定义的 64 位接口 ID，

⊖ EUI-64 指的是由 IEEE 定义的 64 位扩展唯一标识符，本章后面涉及修正的 EUI-64 算法。

RFC 3587 定义的全局单播地址格式并如图 2-5 所示。

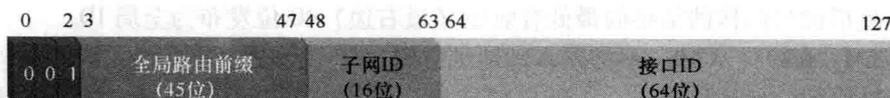



图 2-5　全局单播地址格式[18]

前三位是［001］$_2$ 表明全局单播地址空间。接下来的 45 位是全局路由前缀，紧跟着 16 位子网 ID 和 64 位接口 ID。目前的指导方针要求互联网服务提供商分配/48 网络给它们的客户，从而将全局路由前缀分配给用户。然后，每个用户可以定义多达 65535 个唯一子网，每个子网中还剩下 16 位子网 ID 字段。

### 2.3.3.3　fc00::/7——唯一本地地址空间

在本书参考文献［19］RFC 4193 中定义的唯一本地地址（Unique Local Address，ULA）空间，旨在提供通常在一个站点内的本地分配和路由的 IP 地址。RFC 4193 指出："这些地址没有被期待成为全局因特网路由地址。"因此，并不像 RFC 1918 中严格定义私有的 IPv4 地址空间一样，唯一本地地址空间本质上是私有地址空间，提供"本地"寻址同时仍然具有较高概率保证全局唯一。唯一本地地址空间的格式如图 2-6 所示。

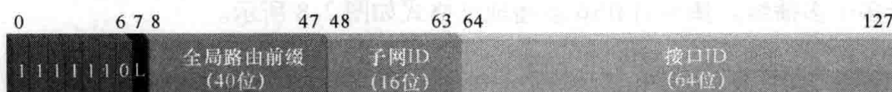



图 2-6　唯一本地地址格式[19]

前 7 位，第 0～6 位是［1111 110］$_2$ = fc00::/7，它标识了唯一本地地址。第 8 位，如果全局 ID 是本地分配的，"L"位设置为"1"。而"L"位设置为"0"标识未定义，虽然互联网社区（如 IETF）讨论通过互联网注册机构启用该设置实现全局唯一的本地地址。40 位全局 ID 字段用来标识全局唯一前缀，它必须通过伪随机算法分配，而不是顺序的。在任一情况下，生成的/48 前缀包括组织的 ULA 地址空间，从而可以使子网分配给内部使用。子网是由一个 16 位的子网 ID 来定义，它的接口 ID 是 64 位字段。

一个利用伪随机方法来得出一个全球独一无二的全局 ID 就如 RFC 4193 中推荐的计算散列（也称哈希，hash）方法㊀如下：

- 网络时间协议（Network Time Protocol，NTP）服务器所报告的 64 位 NTP 格式的当前时间。

---

㊀ 一个散列是将散列化的数据和一个随机值通过运行数学运算而产生的。在这种情况下需要一个特定的数学算法、安全散列算法 1 或 SHA－1。

● 级联完成此算法的主机接口的 EUI-64 接口 ID。

然后散列操作的结果的最低有效位（最右边）40 位发布为全局 ID。

#### 2.3.3.4 fe80::/10——链路本地地址空间

链路本地地址仅用于一个特定的链接，比如一个以太网链接；链路本地目标地址的数据报不会被路由。也就是说，具有链路本地地址的数据报无法到达相应链接之外的链接。这些地址用于地址自动配置和邻居发现，稍后将会展开讨论。链路本地地址格式如图 2-7 所示。

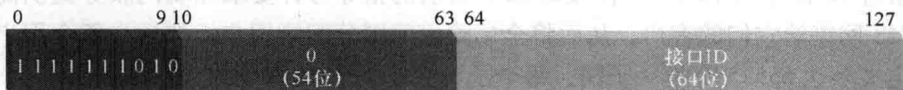



图 2-7 链路本地地址格式[15]

链路本地前缀为 fe80::/10，紧跟着 54 个 0 位和 64 位的接口 ID。

#### 2.3.3.5 ff00::/8——多播地址空间

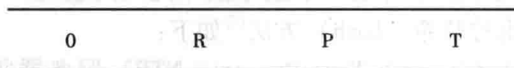
多播地址可以识别在不同节点上一组接口。可以将多播地址想象成是一个范围内的广播。所有多播组成员共享共同的组 ID，因此所有成员会接收到给多播组发送的相同的数据报。一个接口可能有多个多播地址；也就是说，它可能属于多个多播组。基本的 IPv6 多播地址格式如图 2-8 所示。




图 2-8 多播地址格式[15]

前缀 ff00:::/8 标识了这是一个多播地址。紧接着一个字段是一个 4 位的字段称为“标志”。多播地址的格式是由这个标志字段决定的。范围字段指示了多播范围的广度，是否每个节点、链路、全局或是其他范围值，都将在这个字段定义。幸运的是，标志和范围字段可以很容易通过地址中的第 3 个和第 4 个十六进制数字识别出，将在稍后总结。

标志 标志字段是由 4 位组成，下面将会从右到左进行讨论[15]。




● T 位表示由 IANA 分配的多播地址是瞬时状态或是永久状态。T 位是如下这样定义的：

○ T = 0 时——表示多播地址是由 IANA 分配的永久分配（众所周知）多播地址。在这种情况下，接下来的 112 位多播地址是 112 位的组 ID 字段（见

图 2-9）。




图 2-9 标志 T=0 时的多播地址

IANA 到目前为止已经分配了大量的组 ID㊀。例如，组 ID=1 意味着相关范围内的所有节点（由范围字段定义），ID=2 为指定范围内的所有路由器。范围字段如下定义，永久分配的多播地址例子为

· ff01::1=链路上所有节点
· ff02::2=链路上所有路由器
· ff05::1=站点上所有节点
· ff05::2=站点上所有路由器
· ff0e::2=互联网上所有路由器

○T=1——这个多播地址是一个暂时的多播地址。这可以分配给一些特殊的多播会话或是应用程序。Ff12::3:f:10 是其中一个例子。

● P 位用于表示多播地址是否基于单播网络地址前缀。P 位的定义㊁如下：

○P=0——这个多播地址不是根据网络前缀来分配的。当 P=0 时，这个多播报文的格式跟上面描述的一样（当 T=0）也有 112 位组 ID 字段。

○P=1——这个多播地址是根据单播子网地址的网络前缀分配的。单播子网地址拥有这类多播地址的分配权。这使得与单播地址空间相关的多播地址的分配管理更加简单。如果 P=1，那么 T 位也必须为 1。P=1 时的多播地址格式如图 2-10 所示。




图 2-10 P=1 时的多播地址格式[15]

当 P=1，范围字段紧跟着 8 个 0 位（保留），一个 8 位前缀长度字段，一个 64 位网络前缀字段和一个 32 位组 ID 字段。前缀长度字段代表对应单播网络地址的前缀长度。网络前缀字段包含对应单播网络的前缀，组 ID 则是相关多播的组 ID。

例如，如果 2001:db8:b7::/48 的单播前缀被分配给一个子网，那么对应基于单播的多播地址将会为，ff3s:0030:2001:db8:b7::g。这时有：

---

㊀ 最新的分配请参考 http://www.iana.org/assignments/ipv6-multicast-addresses。

㊁ P 位是在本书参考文献［20］RFC3306 中定义的。

· ff = 多播前缀

· 3 = $[0011]_2$，即 P = 1，T = 1

· S = 有效的范围，将会在下一节讨论

· 00 = 保留位

· 30 = 十六进制表示的前缀长度 = $[0011\ 0000]_2$ = 十进制数 48，例子中的前缀长度

· 2001：db8：b7：0 = 2001：0db8：00b7：0000 为在 64 位的网络前缀字段的 48 位网络前缀

· g 为 32 位的组 ID

当前缀长度字段 = FF，s≤2，并且 P = T = 1 时，一种格式的特殊情况将会发生。在这种情况下，并不是由单播网络地址组成网络前缀字段，这个字段将会由各自接口的接口 ID 组成。接口 ID 必须经过重复地址检测，以保证其独一无二，这在本章后面会进行讨论。在这种特殊情况下，范围字段必须是 0、1 或 2，代表着本地接口或本地链路的范围。链路范围内的多播地址格式是由 RFC 4489（即本书参考文献［21］）规定的 IPv6 地址结构扩展来定义的。

在标志字段的 R 位是多播交汇点（Rendezous Point，RP），它可以使多播组潜在的订阅者优先以临时方式而非永久加入该组。如果 R 位设置值为 1，那么 P 位和 T 位也必须设置成 1。当 R = 1，多播地址是基于单播前缀的，但是 RP 接口 ID 也是特定的。这种当 R = 1 时的多播地址格式，大致等同于 R = 0 和 P = 1，不同之处在于保留字段被分成一个 4 位的保留字段交汇点接口 ID（Rendezous Point Interface ID，RIID）字段（见图 2-11）。

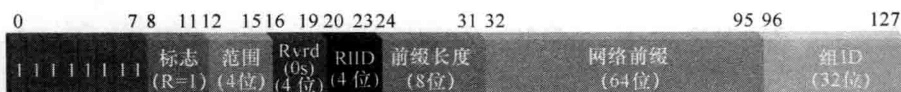



图 2-11 R = 1 时的多播地址

● RP 的 IP 地址通过串联 RIID 字段的值与相应的网络地址前缀确定。例如，如果一个 RP 在［单播］网络时是 2001：db8：b7：：6，那么关联的多播地址会为 ff7s：0630：2001：db8：b7：g。其中，s 为一个有效范围内定义，g 为 32 位的组 ID。

该地址的明确细分如下：

· ff = 多播前缀

· 7 = $[0111]_2$，即 R = 1，P = 1，T = 1

· s = 一个合法的范围

· 0 = 预留位

· 6 = RIID 字段，要被追加到网络前缀字段之后

· 30 = 十六进制表示的前缀长度 = $[00110000]_2$ = 十进制下 48，实例中的前缀长度

· 2001：db8：b7：0 = 2001：0db8：00b7：0000 为 64 位网络前缀字段中的 48 位网络前缀

· g 为 32 位的组 ID

· 第一个标志位是保留位并设为 0。

#### 2.3.3.6　多播标志总结

谁想到多播寻址位这么复杂？但通常情况下，复杂性带来了灵活性！总的来说，按照目前的定义，以上位规定的最终结果产生以下有效标志字段的值。在标志字段之后紧跟着 8 个“1”位，这 8 个紧跟着 4 位标志字段的位为“有效前缀”（见表 2-3）。

表 2-3　多播标志总结

| 标识（二进制） | 有效前缀 | 说　明 |
|---|---|---|
| 0000 | ff00：：/12 | 永久分配 4 比特域范围内的 112 位组 ID |
| 0001 | ff10：：/12 | 暂时分配 4 比特域范围内的 112 位组 ID |
| 0011 | ff30：：/12 | 临时分配基于多播地址的单播前缀 |
| 0111 | ff70：：/12 | 临时分配基于多播地址的拥有接口交汇点接口 ID 的单播前缀 |
| 其他 | — | 未定义 |

**范围**　很自然的，范围字段定义了多播地址的范围或者说定义了多播地址所能通信到的地方。这被多播路径上的路由器用来判定多播能否到达对应的范围。请注意，范围以外的其他本地接口、本地链路和全局地址必须在路由器定义好范围，从而形成通信可达性的约束。

#### 2.3.3.7　特殊的多播地址

**请求节点多播地址**　多播地址的一种，每个节点都必须支持一种形式的请求节点多播地址。这个地址是用来检测重复的地址，并作为邻居发现协议（邻居请求信息）的一部分，用以修复给定主机的链路层地址。请求节点多播地址是在众所周知的前缀 ff02：：1：：ff00：0/104（多播范围 = 链路）后面加上请求节点接口 ID 的低位（最右边）24 位组成。

例如，假设一个节点希望解决设备（接口）的链路层地址，它使用的 IP 地址为 2001：db8：4e：2a：3001：fa81：95d0：2cd1。使用低 24 位、十六进制的 d02cd1，该设备会将请求发送到 ff02：：1：ffd0：2cd1（见图 2-12）。

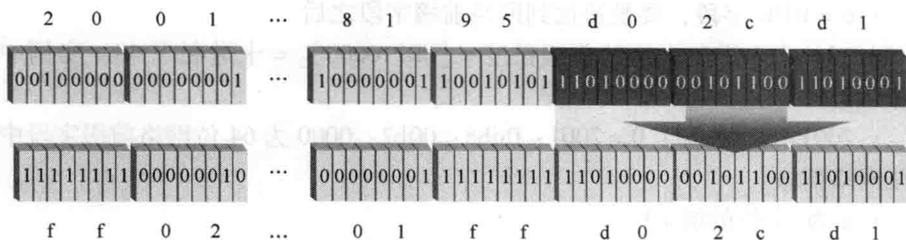



图 2-12 请求节点多播地址的推导[15]

## 2.3.4 IPv6 的因特网控制报文协议

ICMP 是 IPv4 的一部分，它用作回送网络错误、诊断、资源状态给 IP 节点。ICMPv6 是 IPv6 相关的协议，它提供了相似的功能，但它也对 IPv6 中的核心特性，如邻节点发现、移动 IPv6 和多播路由发现，提供了便利。就如下面将讨论的，邻节点发现是一个完整的 IPv6 功能，支持路由器发现、第 2 层地址发现、地址自动配置、重复地址检测和邻居不可达检测。因此，ICMPv6 是 IPv6 不可分割的组成部分，虽然消息被编码为上层协议，但是 ICMPv6 被分配给一个 IPv6 报头“下一个报头”，其代码为 58。ICMPv6 结构格式如图 2-13 所示。

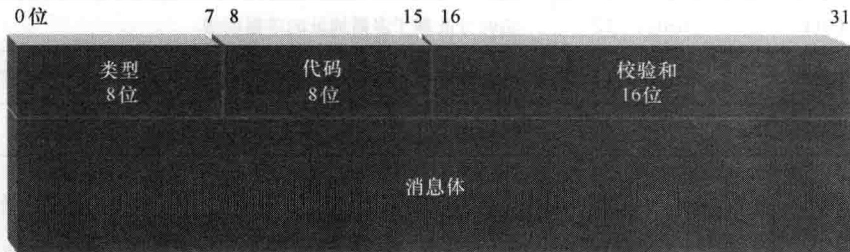



图 2-13 ICMPv6 结构格式

表 2-4 给出了多播范围字段说明，表 2-5 给出了 ICMPv6 类型和代码值[23]。类型值 1 ~ 4 代表错误信息，128 ~ 255 是用作表示信息和诊断信息（见表 2-5）。

表 2-4 多播范围字段说明

| 范围字段 | | 范 围 | 描 述 |
|---|---|---|---|
| 二进制 | 十六进制 | | |
| 0000 | 0 | 保留 | 保留 |
| 0001 | 1 | 本地接口 | 节点上只用于回环传输的接口 |
| 0010 | 2 | 本地连接 | 只有在多播报文传输时连接 |

（续）

| 范围字段 | | 范　围 | 描　述 |
|---|---|---|---|
| 二进制 | 十六进制 | | |
| 0011 | 3 | 保留 | 保留的 |
| 0100 | 4 | 本地管理 | 最小范围的管理配置。这不是基于物理连接或其他多播相关配置 |
| 0101 | 5 | 本地站点 | 限于由管理者定义的站点 |
| 0100、0111 | 6、7 | 未分配 | N/A |
| 1000 | 8 | 本地组织 | 管理者定义的一个组织实体内的多个站点组成 |
| 1001 ~ 1101 | 9 ~ D | 未分配 | N/A |
| 1110 | E | 全局范围 | 无限制 |
| 1111 | F | 保留 | 保留的 |

**表 2-5　ICMPv6 类型和代码值**

| 类　型 | | 代　码 | |
|---|---|---|---|
| 值 | 含　义 | 值 | 含　义 |
| 0 | 保留的 | | |
| 1 | 目标地址不可达 | 0 | 没有到目标地址的路由 |
| | | 1 | 管理策略禁止与目标地址通信 |
| | | 2 | 超出源地址的范围 |
| | | 3 | 地址不可达 |
| | | 4 | 端口不可达 |
| | | 5 | 源地址失败的入口/出口策略 |
| | | 6 | 拒绝路由到目的地址 |
| | | 7 | 源路由头错误 |
| 2 | 报过大 | 0 | |
| 3 | 超时 | 0 | 传输中超过跳数限制 |
| | | 1 | 分片重组超时 |
| 4 | 参数问题 | 0 | 遇到了错误的报头字段 |
| | | 1 | 遇到了不能识别的下一个报头类型 |
| | | 2 | 遇到了不能识别的 IPv6 选项 |
| 5 ~ 99 | 未分配 | | |
| 100、101 | 私有实验 | | |
| 102 ~ 126 | 未分配 | | |

（续）

<table>
<tr><th colspan="2">类　型</th><th colspan="2">代　码</th></tr>
<tr><th>值</th><th>含　义</th><th>值</th><th>含　义</th></tr>
<tr><td>127</td><td>保留给 ICMPv6 错误消息扩展</td><td></td><td></td></tr>
<tr><td>128</td><td>回送请求</td><td>0</td><td></td></tr>
<tr><td>129</td><td>回送应答</td><td>0</td><td></td></tr>
<tr><td>130</td><td>多播监听查询</td><td>0</td><td></td></tr>
<tr><td>131</td><td>多播监听报告</td><td>0</td><td></td></tr>
<tr><td>132</td><td>多播监听完成</td><td>0</td><td></td></tr>
<tr><td>133</td><td>路由器请求</td><td>0</td><td></td></tr>
<tr><td>134</td><td>路由器通告</td><td>0</td><td></td></tr>
<tr><td>135</td><td>邻节点请求</td><td>0</td><td></td></tr>
<tr><td>136</td><td>邻节点通告</td><td>0</td><td></td></tr>
<tr><td>137</td><td>重定向信息</td><td>0</td><td></td></tr>
<tr><td rowspan="3">138</td><td rowspan="3">路由器重编号</td><td>0</td><td>路由器重编号指令</td></tr>
<tr><td>1</td><td>路由器重编号结果</td></tr>
<tr><td>255</td><td>序列编号重置</td></tr>
<tr><td rowspan="3">139</td><td rowspan="3">ICMP 节点信息查询</td><td>0</td><td>数据字段包含 IPv6 地址</td></tr>
<tr><td>1</td><td>数据字段包含一个名称或者是空的(No-Op)</td></tr>
<tr><td>2</td><td>数据字段包含 IPv4 地址</td></tr>
<tr><td rowspan="3">140</td><td rowspan="3">ICMP 节点信息响应</td><td>0</td><td>成功回复——数据字段可能为空，也可能不为空</td></tr>
<tr><td>1</td><td>响应者拒绝回复——数据字段为空</td></tr>
<tr><td>2</td><td>查询的类型对于响应者是未知的——数据字段为空</td></tr>
<tr><td>141</td><td>逆向邻节点发现请求消息</td><td>0</td><td></td></tr>
<tr><td>142</td><td>逆向邻节点发现通告消息</td><td>0</td><td></td></tr>
<tr><td>143</td><td>第 2 版多播监听报告</td><td>0</td><td></td></tr>
<tr><td>144</td><td>本地代理地址发现请求消息</td><td>0</td><td></td></tr>
<tr><td>145</td><td>本地代理地址发现应答消息</td><td>0</td><td></td></tr>
<tr><td>146</td><td>移动前缀请求</td><td>0</td><td></td></tr>
<tr><td>147</td><td>移动前缀通告</td><td>0</td><td></td></tr>
<tr><td>148</td><td>认证路径请求消息</td><td></td><td></td></tr>
<tr><td>149</td><td>认证路径通告消息</td><td></td><td></td></tr>
</table>

（续）

| 类 型 | | 代 码 | |
|---|---|---|---|
| 值 | 含 义 | 值 | 含 义 |
| 150 | 实验移动协议所用的ICMP消息 | | |
| 151 | 多播路由器通告 | | |
| 152 | 多播路由器请求 | | |
| 153 | 多播路由器中止 | | |
| 154 | FMIPv6消息（快速移动IPv6切换） | 0，1 | 保留的 |
| | | 2 | 代理通告的路由器请求（RtSolPr） |
| | | 3 | 代理路由器通告（PrRtAdv） |
| | | 4，5 | 弃用和不可用 |
| 155 | RPL控制消息（暂时分配给“低功率和损耗网络的路由协议） | | |
| 156~199 | 未分配 | | |
| 200~201 | 私有实验 | | |
| 255 | 为ICMPv6通知消息扩展保留 | | |

下面还要仔细介绍IPv6某些与ICMPv6相关的特性。

### 2.3.5 IPv6 Ping命令

就像在IPv4中用的“ping”命令一样，IPv6支持类似的功能。相对于“ping”，通常在命令行中运行“ping6”。此功能利用ICMPv6类型128作为ping6命令的接口，或称为“回送请求”，而ICMPv6类型129则作为“回送应答”信息。

### 2.3.6 多播侦听发现

多播侦听发现（Multicast Listener Discovery，MLD）功能使得路由器可以找到直连链路上存在的多播侦听器，也就是被配置了要侦听多播地址的节点及相应的多播地址。多播主机和地址的确定，可以使路由器去路由和发送相应的多播IPv6数据报。当在给定链路上发送MLD查询时，路由器使用本地链路IPv6地址作为其源IPv6地址。MLD的第1个版本，即MLDv1，是RFC 2710（即本书参考文献［24］）定义的。第2个版本，MLDv2[25]与MLDv1是可兼容的，它增加了源过滤器。也就是说，它允许一个多播侦听器只接收从特定的源地址中发出的多播数据报。下面给出4个MLD信息类型定义，每一个都有对应的ICMPv6类型值：

- 多播侦听查询（类型值为 130）——允许路由器去查找侦听器。
  - 一般查询——用于路由器去了解哪一个多播地址拥有侦听器。
  - 特定多播地址查询——用于确定一个多播地址是否拥有侦听器。
  - 多播地址和特定源查询——只在 MLDv2 可用，确定一个给定源的特定多播地址是否拥有侦听器。
- 多播侦听报告（类型值为 131）——可以使一个多播侦听器回复路由查询，以表明它希望接收这样寻址的多播地址数据报。
- 多播侦听完成（类型值为 132）——允许一个多播侦听器去声明它已经完成对给定多播地址传输的侦听。这个信息会被发送到链路范围内的所有路由器中的多播地址，ff02::2。
- MLDv2 多播侦听报告（类型值为 143）——由 MLDv2 主机通知连接到链路上的路由器它的多播侦听状态，如它是否正在侦听某个地址。

### 2.3.7 多播路由发现

多播路由发现（Multicast Router Discovery，MRD）过程被定义来识别多播路由中涉及的路由器，以及允许交换机确定哪些交换端口应该在多播通信中被桥接。因此，MRD 报文目的是将本地链路的跳数限制为“1”。MRD 为 IPv6 提供了三种 ICMPv6 消息类型：

- 多播路由器通告（ICMPv6 类型值为 151）——路由器周期性通告或回复这个信息给所有的侦听者多播地址（ff02:: 6a），通告其参与多播转发。
- 多播路由器请求（ICMPv6 类型值为 152）——可以使节点去请求多播路由器发送多播路由器通告。
- 多播路由器终止（ICMPv6 类型值为 153）——由路由器发送来表明不再参与多播信息转发了。

### 2.3.8 邻节点发现协议

邻节点发现协议（Neighbor Discovery Protocol，NDP）在 IPv6 中提供了相当重要的网络操作功能。

- NDP 允许一个节点去发现链路上的路由器和链路上的 IPv6 前缀配置。
- NDP 提供了重定向功能去指示一个节点到一个更好的首跳路由器（或本地连接的主机）。
- 无状态地址自动配置（Stateless Address Autoconfiguration，SLAAC）利用 NDP 给 IPv6 主机地址进行自动匹配。
- 重复地址检测（Duplicate Address Detection，DAD）允许节点判断自己准备使用的地址是否已被子网的其他节点使用。

● 一个使用 NDP 的节点进行地址解析，找出子网的其他 IPv6 节点，并确定它们的链路层地址。

● NDP 支持节点不可达检测，适用于邻居节点的不可达检测，也就是在同一链路上的不可达节点。

路由器发现使得 IPv6 节点能够自动识别子网上的路由器，使其不需要手动配置设备的 IP 配置内的默认网关。路由器发现使得一个设备可以识别分配给链路的网络前缀和对应的前缀长度，当然也包括其他参数，如可用的地址分配和域名服务。对于地址自动配置，这个信息是必不可少的。

路由器发现过程需要每个路由器周期性地发送通告给每个配置子网以显示，它的 IP 地址，它能够提供的默认网关功能，它的链路层地址，链路上的网络前缀（包括对应的前缀长度和可用地址生命周期），以及其他的配置参数。

路由器通告还能指出 DHCPv6 服务器是否可用于地址分配或其他配置。路由器通告中的 M 位（管理地址配置标志）表明 DHCPv6 服务是可用于地址和配置的设置的。O 位（其他配置标志）则表明通过 DHCPv6 获取 IP 地址以外的其他配置参数，如可能会包含链路上的设备可以查询哪些 NTP 服务器。关于 M 和 O 位的解释见表 2-6。

**表 2-6 路由器通告中的 M 和 O 位**

| 标　志 | O 位 = 0 | O 位 = 1 |
|---|---|---|
| M = 0 | 没有 DHCPv6 服务可用 | DHCPv6 服务仅可用作配置信息，不包含地址分配 |
| M = 1 | DHCPv6 服务可用于地址分配和配置信息 | DHCPv6 服务可用于地址分配和配置信息 |

节点也可以通过使用路由器请求信息来请求路由器通告，只要发给本地链路路由器多播地址（ff02::2）即可。ICMPv6 每个邻节点发现信息的类型代码总结如下：

● 路由器请求（ICMPv6 类型值为 133）——允许一个主机立即请求路由器产生路由通告。

● 路由器通告（ICMPv6 类型值为 134）——路由器用这个信息类型去定期地通告它们的存在和可用的前缀信息或是响应路由器请求信息。

● 邻节点请求（ICMPv6 类型值为 135）——可以使一个主机去确定一个链路上的邻节点的链路层地址，并且进行重复地址检测。

● 邻节点通告（ICMPv6 类型值为 136）——由主机发送用于响应邻节点请求或链路层地址变化。

● 重定向（ICMPv6 类型值为 137）——由路由器发送去通知链路主机一个给定目标地址的更好的首跳方式。更好的首跳，可能是另一个路由器或者是另

一台链路上的主机，如它的地址有同样的前缀没有在链路上通告。

### 2.3.9 安全邻节点发现

安全邻节点发现（Secure Neighbor Discovery，SEND）协议在邻节点发现信息中添加了数字签名，以此减少篡改响应的风险，特别是在 Ad Hoc 网络上。使用 SEND 的主机必须配置信任锚和一对公钥私钥，这可以使主机保证是它网络上的路由器。节点使用 SEND 将加密生成地址（Cryptographically Generated Address，CGA）附加为 NDP 消息中的选项。CGA 是由主机的公钥和私钥推导而出的。一个 64 位散列值的公钥（还有一些辅助的参数）和一个 64 位的子网前缀被用于生成 CGA。那对应的私钥也可以被用作 NDP 消息的签名，并且对应的数字签名已经在 RSA 签名 NDP 消息中被添加。虽然攻击者可以用自己推导的一对公钥、私钥来生成一个 SEND 信息，但是它不能模拟自己是子网中的某个节点。

通过路由器证书上的公共密钥与主机上配置的公共密钥对比，路由器的身份被进一步验证（信任锚）。以下两个额外的 ICMPv6 消息类型就是为此设定的：

- 认证路径请求（ICMPv6 类型值为 148）——主机发送这条消息给路由，请求获得主机配置的信任锚中的认证路径。
- 认证路径通告（ICMPv6 类型值为 149）——路由器发送这条消息作为认证路径请求的响应，提供路由器的认证和/或 DNS 的完全合格域名（Full Qualified Domain Name，FQDN）或者 X. 501 名形式的信任锚。

### 2.3.10 逆向邻节点发现

逆向邻节点发现（Inverse Neighbor Discovery，IND）定义了解决 IPv6 地址与一个已知链路层地址关联的过程。与“逆向地址解析协议（Address Resolution Protocol，ARP）”的概念相似，IND 允许第二层网络解析第三层信息，如帧中继网络。IND 也是通过 ICMPv6 实现的：

- IND 请求（ICMPv6 类型值为 141）——请求提供的目标链路层地址所对应的 IPv6 地址。
- IND 通告（ICMPv6 类型值为 142）——响应 IND 请求，返回目标链路层地址确定的一个或多个 IPv6 地址。

### 2.3.11 路由器重编号

路由器重编号概念的提出，旨在使重编号网络像单独的主机地址自动配置一样简单和自动地进行。路由重编号在 RFC 2894（即本书参考文献［26］）中进行了定义，它使用了 ICMPv6 消息（ICMPv6 类型值为 138）的“前缀控制操

作”去通知一个路由器（当以多播方式发送时，可以是多个路由器）去增加IPv6 前缀或者删除改变一个已被配置的前缀。前缀信息包括前缀地址、长度，以及与地址相关的生命周期值和标志。鉴于远程网络重编号的关键性，路由器要求采用安全策略来进行路由器重编号信息，去证明发送者是经过认证的，来检查信息的完整性。消息队列编号也用于防止重复攻击。

## 2.3.12　节点信息查询

节点信息查询消息允许从 IPv6 主机中请求主机名、IPv6 和 IPv4 地址信息。如果你觉得这听起来跟 DNS 提供的服务有点像，那么你对了。然而，根据 RFC 4620（即本书参考文献［27］），这种模式的解决方案是“目前仅限于诊断工具、调试工具和网络管理”的。与查询 DNS 消息不同的是，查询是指被发送到节点的信息查询地址。

节点信息查询，是使用ICMPv6 发送并被赋予一个正常的IPv6 目标地址，或是使用节点信息查询多播地址。用这种链路范围多播地址格式使得一个 IPv6 地址只根据接收者的主机名组成；如果 IPv6 地址已经被知道并且主机名信息被请求，那么 IPv6 地址可能会被用于目标地址。当 IP 地址信息被一个已知主机名请求时，对主机名做128 位 MD-5 算法散列运算，结果的头 24 位被加到 ff02：：2：ff00：0/104 的前缀。每个节点接收到给这个节点信息查询地址的消息，并将其地址的最后24 位与自己主机名散列值的头24 位比较；如果匹配，那么接收者会回复请求信息。

节点信息查询消息如下：

- 节点信息查询（ICMPv6 类型值为139）——允许一个节点提供IPv6 地址、IPv4 地址或主机名，来查询关于另一个节点的信息。
- 节点信息响应（ICMPv6 类型值为 140）——允许一个节点响应请求信息或拒绝请求（见图2-14）

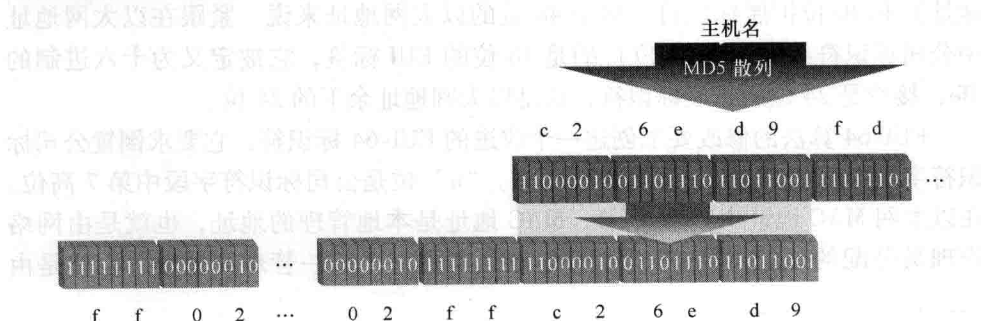



图 2－14　请求节点信息查询地址

## 2.4 IPv6 地址自动配置

IPv6 宣扬的主要好处之一，是其使设备能够自动配置它们自己的 IPv6 地址，这些地址是独一无二且与它现在正在相连的子网相关[⊖]。这种行为是受在本书 NDP 部分讨论的路由器通告管理位（M）和其他位（O）来进行管理的。现在有三种基本形式的 IPv6 地址自动配置：

**无状态。**这个形式是“无状态”的，因为它是不依赖状态或外部分配机制的，如 IPv6 动态主机配置协议（DHCPv6）。设备试图在免除外界和用户的干扰下去配置它自己的 IPv6 地址。

**有状态。**有状态的形式只依赖一个外部地址分配机制，如 DHCPv6。DHCPv6 服务器分配 128 位 IPv6 地址的方式类似 DHCP 对 IPv4 的操作。

**有状态与无状态组合。**这一过程包括无状态地址自动配置与有状态配置涉及的附加 IP 参数的组合。这通常需要先使用无状态方法，然后利用 DHCPv6 获取额外的参数或选项，如网络时间协议服务器去查询在给定网络上的时间。

在最基本的层面，IPv6 地址的无状态自动配置包括串联设备连接到网络的地址（你在哪里）和设备的接口 ID（你是谁）。那么，首先考虑设备如何确定连接到的是哪一个网络地址。

### 2.4.1 改进的 EUI-64 接口标识符

一旦一个节点识别它所连接的子网，它可能通过自己的接口 ID 来完成地址自动配置。根据 IPv6 寻址体系结构规定，所有单播 IPv6 地址，除了以二进制 $[000]_2$ 开头的，必须根据 64 位接口 ID 由改进的 EUI-64 算法推导生成。“未改进的”EUI-64 算法要求串联 24 位由 IEEE 发布给每个网络接口的硬件制造商的组织唯一标识符（Organizational Unique Identifier，OUI）（如最初的 24 位以太网地址）和 40 位扩展标识符。对于 48 位的以太网地址来说，紧跟在以太网地址中公司标识符部分（头 24 位）的是 16 位的 EUI 标签，它被定义为十六进制的 fffe，接着是 24 位的扩展标识符，这是以太网地址余下的 24 位。

EUI-64 算法的修改要求创建一个改进的 EUI-64 标识符，它要求倒置公司标识符字段中的“u”位（全局/本地位）。“u”位是公司标识符字段中第 7 高位。在以太网 MAC 地址中，当 u = 1，MAC 地址是本地管理的地址，也就是由网络管理员分配的地址；当 u = 0，MAC 地址就是一个统一管理的地址，也就是由

⊖ 注意一些 IPv4 的协议栈，如 Microsoft Windows 2000 和 XP，用 IPv4 的本地链路地址空间 169.254.0.0/16 完成地址的自动配置。

NIC 制造商从 IEEE 中分配的。倒置“u”位的目的[15]是使它更容易让网络管理员手动配置接口 ID，使其增量计数，如用：：1、：：2、：：3，等来代替：：200：0：0：1、：：200：0：0：2、：：200：0：0：3 等。这些地址是被要求用来表明该地址是本地管理的（u=1）。因此，该算法对于一个 48 位的 MAC 地址的处理是倒置“u”位，并且在公司标识符和接口标识符之间插入十六进制数的 fffe。这一点在下面的例子中说明，使用 MAC 地址 AC-62-E8-49-5F-62，由此产生的接口 ID 是 ae62：e8ff：fe49：5f62（见图 2-15）。

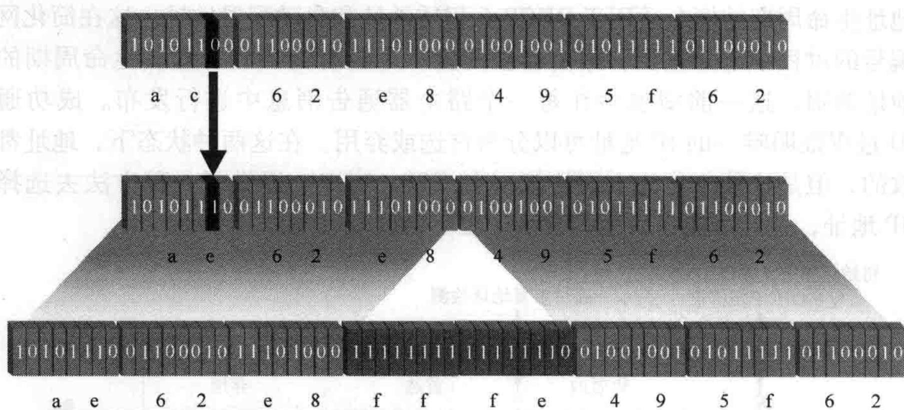



图 2-15　修改 EUI-64 接口 ID 例子[28]

对于非以太网 MAC 地址，该算法要求是的链路层地址作为接口 ID，从左开始用 0 填充。对于没有链路层可用的情况，如在一个拨号链路、一个唯一标识符使用另一种接口地址的、一个序列号或者其他设备独有的标识符，也是推荐的。

然而改进的 EUI-64 算法简化了自动配置的过程，它也创建一个静态的具有“跟踪”设备的能力的接口 ID。它可以简单地识别一个已知的 MAC 地址的设备。RFC 4741（即本书参考文献［29］）定义了如何进行私有扩展来进行随机接口 ID 的推导和变换来解决相关问题。

接口 ID 可能不是唯一的，特别是并非由唯一的 48 位 MAC 地址推导出的接口 ID。因此设备在确认使用新地址前，必须执行 DAD。完成 DAD 的过程之前，设备地址会被认为是暂时的。

## 2.4.2　重复地址检测

DAD 使用 NDP，它需要设备发送一个 IPv6 邻节点请求数据报去它刚刚推导出的 IPv6 地址（或者从 DHCPv6 获得的）来确认此 IP 地址是已经被占用。在短暂的延迟后，设备也发送一个邻节点请求数据报到与请求地址相对应的多播

地址。

如果另一个设备已经使用了此 IP 地址，它会响应一个邻节点通告数据报，其自动配置过程会停止。也就是说，需要手动介入，使用可选的接口 ID 对设备进行配置。如果邻节点通告数据报没有被接收，那么此设备可以假设这个地址是唯一的，并把它配置到相应的接口上。参与邻节点请求和通告的过程，不仅对地址自动配置有用，也对静态配置的地址和从 DHCPv6 获取的地址有用。

有效的 IPv6 地址是有生命周期的。在一些情况下，生命周期是无穷的，但是地址生命周期的概念适用于 DHCPv6 租赁地址和自动配置地址。这在简化网络重编号的过程中十分重要。路由器被配置一个具有恰当和有效的生命周期的网络地址前缀，这一前缀也会在每一个路由器通告消息中进行发布。成功通过 DAD 过程证明唯一的 IP 地址可以分为首选或弃用。在这两种状态下，地址都是有效的，但是这种分化为上层协议（如 TCP、UDP）提供了一种方法去选择一个 IP 地址，并且可能在后续的会话中也不会改变（见图 2-16）。

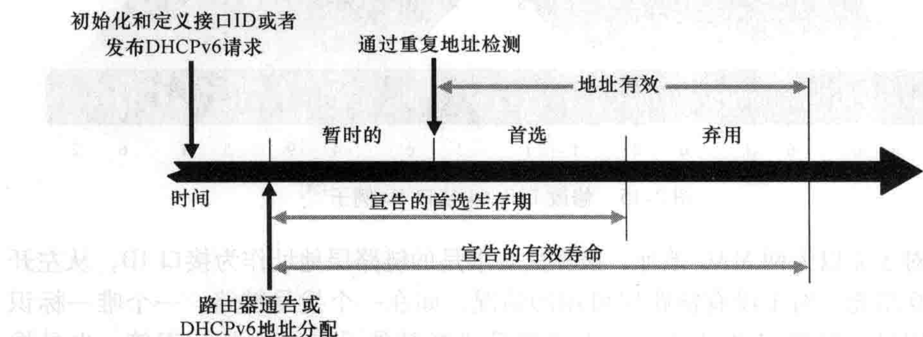



图 2-16 IPv6 地址生命周期（源于本书参考文献[30]）

设备通过发送路由器通告消息来刷新选定有效时间值。当一个地址前缀的选定时间值到期，那么与之相关联的地址也会被弃用，尽管它仍然是有效的。因此，已过时的状态具有一个过渡期，在这期间的地址仍然具有相应功能，但是不能用来发起新的通信。一旦地址的有效生命周期结束，那么这个地址也就不再有效。应该为子网重新分配一个不同的网络前缀，这样路由器可以配置来发布这个新前缀，网络上的设备会用新的前缀来代替过期的旧前缀进行自动配置。

## 2.5 移动 IPv6

IPv6 的移动性支持，或者称之为移动 IPv6，使得 IPv6 节点从链路到链路移动时无缝沟通。这意味着，尽管底层网络传输、数据链路层和物理层网络不断

变化，但上层传输和应用层通信仍旧保持完好。当然改变链路时，如当从 4G 无线服务中移动到 WiFi 网络时，这意味着 IPv6 前缀的改变，那么这个移动设备就必须改变它的 IPv6 地址了。这种连接在当前网络上的 IPv6 地址变化，被称作转交地址。

每一个移动 IPv6 设备也有一个“固定的” IPv6 地址被称为其归属地址。如果这个设备并没有漫游而是“本地”状态的，那么 IPv6 流量路由会很自然地使用它的归属地址。当设备漫游时，它得到一个转交地址，这个有状态或无状态的地址是根据当前所在位置和所连接的网络来自动配置的。然后，移动节点以其归属代理注册转交地址，一个移动 IPv6 配置路由器在链路上为这些移动节点的归属地址进行服务。当移动主机是归属状态，那么归属代理就跟普通的路由器一样给设备路由 IPv6 数据报，如图 2-17 所示。当移动主机在漫游状态，那么归属代理会拦截发给移动主机归属地址的 IPv6 数据报并使用转交地址通过隧道转发给移动主机。

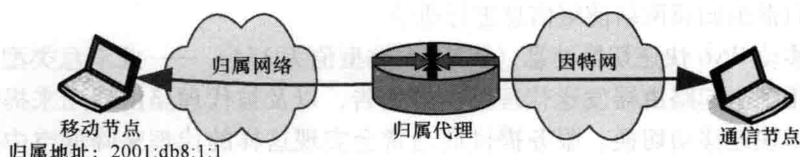



图 2-17　归属移动的移动 IPv6

移动主机和其他主机（通信节点）之间的通信，仍可能以下面两种方式之一进行。使用隧道的方法如上所述，IPv6 数据报可能会直接与归属地址进行通信，这样它们会被归属代理拦截并通过隧道转发给移动主机；返回流量将遵循同样的路线，通过归属代理到达相应的节点。这一过程如图 2-18 左图所示。

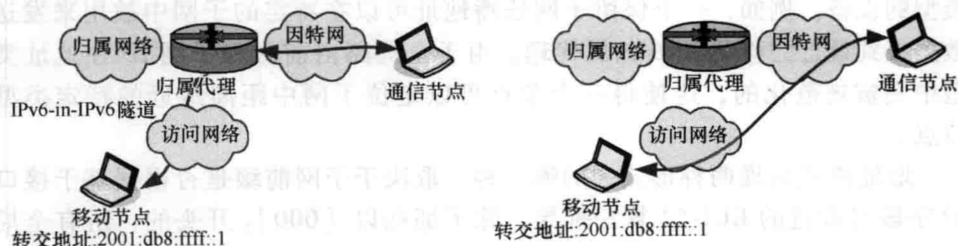



图 2-18　漫游中的移动 IPv6

一般来说，这种“三角路由”的过程不仅低效，还可能导致归属代理资源超载。移动节点和通信节点之间一种更高效更直接的通信方式，如图 2-18 右图所示。这种更有效的路由选择只要求对应节点支持 IPv6，包括移动性扩展报头即可。移动头是用来在移动节点和通信节点携带消息，并用此消息去证实转交

节点和归属地址关联，路由可达性及当移动主机持续移动时通信连接的实时更新。本书第 6 章再详细讨论这部分内容。

作为扩展头的补充，移动 IPv6 利用以下五种 ICMPv6 消息类型：

• 归属代理地址发现请求（ICMPv6 类型值为 144）——允许一个移动节点初始化动态归属代理发现。在移动归属地址前缀附上已知的归属代理任播地址（本书第 3 章会讨论），这允许移动设备识别网络上的归属代理，如漫游时一个归属代理被重新配置了。

• 归属代理地址发现响应（ICMPv6 类型值为 145）——响应归属代理发送的归属代理地址发现请求，以标识它作为归属代理的单播地址。

• 移动前缀请求（ICMPv6 类型值为 146）——允许一个移动设备去收集关于它的归属网络的前缀信息，如归属网络重新配置时归属网络前缀可能发生变化。

• 移动前缀通告（ICMPv6 类型值为 147）——归属代理可以通过使用这个消息与目前的归属网络前缀信息进行通信。

• 移动 IPv6 快速切换消息（ICMPv6 类型值为 154）——此消息类型，被移动节点用来引起路由器发送代理路由器通告，以及被代理路由器用来提供这样的通告给快速移动切换。服务提供商通常会实现这样的代理来降低空中接口开销和提高网络效率。

## 2.6 保留子网任播地址

RFC 2526（即本书参考文献［31］）定义了保留子网任播地址的格式。这些地址被 IPv6 设备用来在一个特定的子网中将数据报路由到距离最近的特定类型的设备。例如，一个保留子网任播地址可以在特定的子网中被用来发送数据报到最近的移动 IPv6 归属代理。由于全局路由前缀和子网 ID 在地址类型中是被规范化的，这使得一个节点可以定位子网中距离最近的特定类型节点。

地址格式呈现两种形式中的哪一种，取决于子网前缀是否根据基于接口 ID 字段经改进的 EUI-64 算法推导。除了那些以 $[000]_2$ 开头的，所有全局单播地址都必须使用 64 位接口 ID 机制，它以接口的链路层地址和之前提到的改进的 EUI-64 算法为推导基础。

1. 如果需要 EUI-64 算法，保留子网任播地址则以下面几个字段串联制定：

◆ 64 位全局路由前缀和子网 ID。

◆ 除了第 7 位是 0，57 位连续的全 1（第 71 位开始从左向右数）。当使用 EUI-64 算法时，这个第 7 位与硬件地址中的公司标识符字段中的“u”位一致。

这个二进制位通常为 0，在这个特定的情况下其代表“归属”。

◆ 七位任播 ID。RFC 2526 定义了一个单一的十六进制数 7e 作为移动 IPv6 归属代理任播地址，其他任播 ID 是保留的。尽管 IANA 可能会根据未来的 IETF RFC 标准分配另外的任播 ID（见图 2-19）。

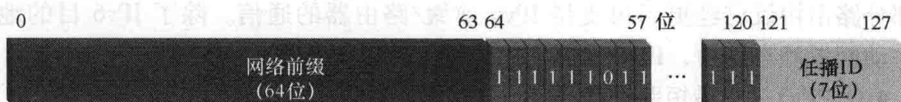



图 2-19 当要求 EUI-64 时保留子网任播地址格式[31]

2. 如果没有使用 EUI-64，而基于全局路由前缀和子网 ID，那么网络前缀长度是任意的 n 位，紧接着是 121 - n 的 1 位，后跟着 7 位任播 ID（见图 2-20）。




图 2-20 当不要求 EUI-64 时保留子网任播地址格式[31]

## 2.7 要求的主机 IPv6 地址

RFC 4284（即本书参考文献［32］）总结了 IPv6 节点、一台实现了 IPv6 的设备和 IPv6 路由器的需求。根据所需的地址，所有 IPv6 节点必须能够识别的自己的 IPv6 地址如下：

- 回环地址（:: 1）。
- 本地链路单播地址（fe80:: <接口 ID>通过自动配置进行配置）。
- 全部节点多播地址（ff0s:: 1，其中 s 为范围）。
- 单播和任播地址在各个接口上自动或手动配置。
- 每个单播和任播地址的请求节点多播地址。
- 节点所属的每个多播组的多播地址。

此外，一个路由器节点还需要以下地址：

- 子网路由器任播地址（<子网前缀>:: /128，子网 ID = 0s）除了在/127 上的点对点路由器链路。
- 所有路由器多播地址（ff0s:: 2，其中 s 为范围）。
- 在路由器中配置的任播地址。

其他设备类型，如 DHCP 和 DNS 服务器，必须认得范围内由 IANA 分配的组 ID（如标志 = 0）对应的多播地址。

## 2.8 IPv6 路由

IPv6 动态路由操作与 IPv4 路由的相同，它使用最长前缀匹配算法。目前，大部分路由协议已经更新可支持 IPv6 前缀/路由器的通信。除了 IPv6 目的地址的手动配置静态路由，IPv6 内部网关协议支持以下几项：

- OSPFv3——最短路径优先。
- 整合 IS-ISv6——中间系统。
- RIPng——下一代路由信息协议。
- 思科 EIGRP——增强内部网关路由协议。
- 边界网关协议（Border Gateway Protocol，BGP）——各种现存的外部网关协议支持 IPv4 以上的版本，也支持 IPv6 协议。

# 第3章　IPv4/IPv6 共存技术

每个考虑部署 IPv6 的组织机构都必须将 IPv4 和 IPv6 一起考虑进去。这是毋庸置疑的，就算对于那些打算部署独立 IPv6 的组织来说也同样如此。这就又说回到了本书第 1 章提到的全局因特网越来越混合的问题。假设一个组织机构要接入到这个无所不在的因特网，那么它们的部署应该同时支持 IPv4 和 IPv6 一段时间。

由于 IPv4 和 IPv6 的兼容性不太友好，互联网社区很早地意识到了 IPv4 和 IPv6 之间互通的必要性。事实上，IETF 至今已经提出超过 20 种 IPv4 和 IPv6 互通的方案，显然并不缺可行方案。但困难的是，在这么多的选择下，如何筛选最适用你的网络的技术。本章，将把重点放在这众多相对成型的 IPv4/IPv6 共存技术上，分别讨论它们的显著特点和优势，从而帮助你作出最适合部署你的网络的选择。

总体来看，这些 IPv4/IPv6 共存技术可大致分为以下这三类：

- 双栈。IPv4 和 IPv6 都在网络设备上获得支持。
- 隧道。将 IPv6 的数据报封装在 IPv4 的数据报中，使用 IPv4 网络进行传输，反之亦然。
- 转换。指 IP 头、IP 地址和端口的转换，如主机、网关或 NAT 设备上实现的那样。

一些设备提供方提出了双重协议策略，会选取本章后续讨论的多种技术进行结合。企业部署可能同样需要多种技术的结合实施，以适应分阶段部署和合作伙伴网络的需要等。而双栈则是设想中最普遍的实现方法。

## 3.1　双栈

双栈方法是 IPv4 和 IPv6 的协议栈在网络设备上的共同实现。这里的设备包括了路由器和其他的基础设备、应用服务器、终端用户设备，它们都需要支持接入两种协议网络层的技术。这些设备都需要被设置 IPv4 和 IPv6 两种协议地址，也可以通过管理员授权对应不同协议的不同途径获得这些地址。

使用双栈进行部署，相比单 IP 的协议栈，根据双栈共用的协议的不同，实施的形式也不尽相同。理想的情况是，只改用双网络层，而应用层、传输层和数据链路层仍然使用共同的协议。这种方式已经在微软 Windows Vista 和

Windows 7 系统上采用，相反微软 Windows XP 对传输层和网络层均采用了双栈结构，而这种方式有时候会导致在双协议栈上的冗余配置。而其他有一些方法，则将双栈的范围扩展到物理层，这就要求网络为 IPv6 和 IPv4 分别提供网络层接口。这种部署方法虽然无法体现分层网络协议模型优势，却是有针对目的的。尤其是对于装载许多复杂的应用和服务的网络服务器，其中的某些应用或服务只支持某一种 IP 时，会变得尤为合理。

### 3.1.1 双栈的实施

双栈设备的部署使用同一物理网络接口。也就是说，IPv4 和 IPv6 在同一物理链路上工作，该双栈网络原理（物理和逻辑）如图 3-1 所示。根据分层协议特性，以太网和其他处于模型第二层的技术均能够为 IPv4 和 IPv6 中的其中一种协议提供负载支持。双栈设备需要同一路由器支持同为双栈的链路，或者 IPv4 和 IPv6 的路由器之间有直接链路连通。为给定的设备分配 IPv4 和 IPv6 地址时，采用以下手段：

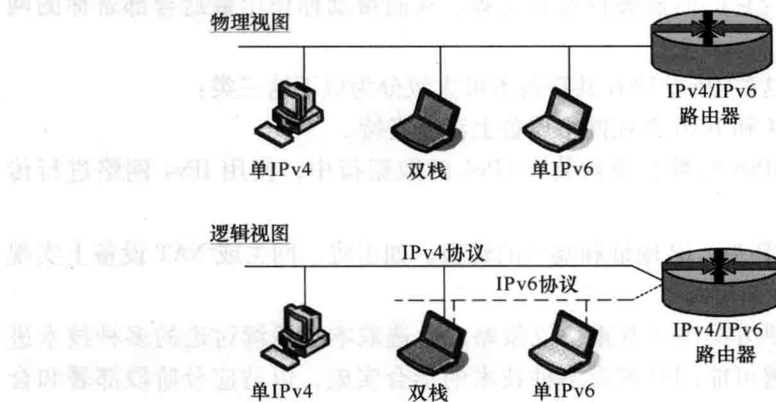



图 3-1　双栈网络原理[28]

- 对于路由器、交换机和接入互联网的设备（如网站、邮件服务器）这样的基础设备，可以人工直接在设备上配置 IPv4 和 IPv6 地址，使得这些设备具有确定的持续一致的地址。当更改这些设备上的地址时，相对应的 DNS 资源记录（A 或 AAAA）也要随之改变。
- 对于其他的非基础设备，除了同样采用静态分配方法外，也可使用动态地址分配。通常使用 DHCP 对 IPv4 地址进行动态分配。而对于 IPv6 地址，对应地可使用 DHCPv6 进行分配，或者使用 SLAAC 进行分配，也可以是两者的结合。注意，IPv4 地址和 IPv6 地址的分配是相对独立的，并不存在通过一次操作就既分配 IPv4 地址、又分配 IPv6 地址的 DHCP；对 IPv4 和 IPv6 两者也可以分别采用

不同的分配策略，如一种用静态分配方法而另外一种采用 DHCP。

路由器被普遍预期会是第一种需要升级支持两种协议的设备。而一个关于信息类的标准——RFC 4554（即本书参考文献［33］）描述了一种无需立即升级路由器，只使用 VLAN 进行配置的全新方法。此方法利用了 VLAN 标识，使二层的交换机可以向一台或多台支持 IPv6 的路由器广播或发送嵌入了 IPv6 负载的以太网帧。将某台路由器（如连接 IPv6 网络（因特网）的网关）升级至支持 IPv6 后，与该路由器连接的交换机端口等都可配置为“IPv6 VLAN”。而其他 IPv6 或双栈设备则可以配置成此 IPv6 VLAN 的成员，类似的，可以配置多个这样的 VLAN。图 3-2 所示的使用 VLAN 的双栈部署，是这种部署的一个示例。

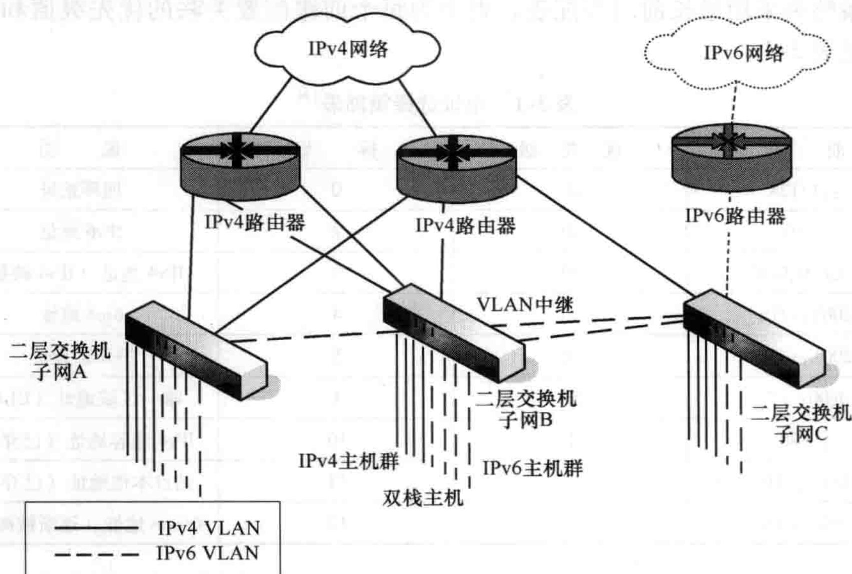



图 3-2 使用 VLAN 的双栈部署[33]

## 3.1.2 使用哪种地址

为选定的设备配置好 IPv4 和 IPv6 地址后，要如何指定何时用何种协议呢？RFC 6724（即本书参考文献［34］）描述了当上层应用⊖没有指定选用 IPv6 设备时，分别如何选择源地址和目的地址的算法。这种情况下，它通常会调用 getaddrinfo（）套接字接口（sockets API），向它的 TCP/IP 栈获取一个目的 IP 地

⊖ 例如，输入一个 IP 地址打开 FTP 会话时，这个目的 IP 地址即被使用（假设目的地址的设备使用的协议版本与源地址一致）。

址。因此，如果你的笔记本电脑是双栈的，当你访问一个“www”网址时，浏览器会使用 getaddrinfo（）调用获取一个或者一组目的 IP 地址。RFC 6724 所定义的这个地址选择算法是通过，使用 getaddrinfo（）调用，将返回应用的 IP 地址按优先级进行排序，从而选择所需的目的地址。而源地址通常是由网络层（如根据你的笔记本电脑上的设置）进行选择，用于初始化选择目的地址的连接。

地址选择基于以下几项输入参数：设备配置的地址、这些地址的状况（如优先与弃用）、地址的范围（如 ULA 与公共地址），以及 A 型或 AAAA 型 DNS 域名解析返回的地址。其过程是，通过设备的 TCP/IP 栈（通常搭建于设备的操作系统上），根据策略表，从候选的地址中，选择优先级最高的或者最适合的地址。策略表采用最长前缀匹配表，表中为每个前缀配置关联的优先级值和标记值，见表 3-1。

**表 3-1 地址选择策略表[34]**

| 前缀 | 优先级 | 标记 | 说明 |
| --- | --- | --- | --- |
| ::1/128 | 60 | 0 | 回环地址 |
| ::/0 | 40 | 2 | IPv6 地址 |
| ::ffff:0:0/96 | 30 | 3 | IPv4 地址（IPv4 映射） |
| 2002::/16 | 20 | 4 | 6to4 地址 |
| 2001::/32 | 10 | 5 | Teredo 地址 |
| fc00::/7 | 50 | 1 | 唯一区域地址（ULA） |
| ::/96 | 1 | 10 | IPv4 兼容地址（已弃用） |
| fec0::/10 | 1 | 11 | 站点本地地址（已弃用） |
| 3ffe::/16 | 1 | 12 | 6bone 地址（逐渐被淘汰） |

DNS 返回的目的地址参照它们各自优先级的值进行排序；优先级越高，在返回的地址列表就越靠前。而同样的，源地址也根据优先级的值划分优先顺序，不过带有目的地址标记值的源地址会更加优先。因此，如果 DNS 域名解析到了一个 ULA 地址，而源设备就是带对应前缀的 ULA 地址，那么 IPv6 头中对应的源地址和目的地址将会被设定为这些 ULA 地址。如果上面给出的策略表按照优先级值的递减顺序排列，那么大致将得到的是如下优先级排列的顺序，依次为回环地址、ULA 地址、IPv6 地址、IPv4 地址、6to4 地址、Toredo 地址。剩下三项的优先级值均为“1”，即不提倡使用。从 IPv4/IPv6 共存的角度看，根据这个默认策略表，显然 IPv6 要优先于 IPv4 被使用。

RFC 6555（即本书参考文献［35］）定义的“快乐眼球双栈（happy eyeballs dual stack）”是影响 IPv4 和 IPv6 连接选择的另一个因素。建立一个连接（由 TCP 或应用创建）的常规步骤是，对上述地址选择返回的地址依次进行连接

建立。“快乐的眼球”方法力图减少潜在的连接延迟，当一个刚建立的 IPv6 连接断开时，紧随其后的 IPv4 连接可以成功建立。实现过程是，假设 IPv6 地址在地址选择中最优先，则先试图建立 IPv6 连接；然后，假设在地址选择过程中返回最靠前的为 IPv4 地址，则在短暂的 300 ms 延迟之后，初始化 IPv4 连接。使用这样的方法，IPv6 连接如果建立失败，那么转为建立 IPv4 连接的延迟就被最小化了，应用的体验受到的影响很小。这个方法的使用前提是，尝试建立的连接所给定的主机域名（如 www. ipamworld. com）已经以 A 型和 AAAA 型存于 DNS 资源记录。因此，建议[㊀]在 DNS 记录中同时使用两种类型保存双栈的域名，而不是只使用“ipv6”这样的标识（如 ipv6. ipanworld. com）。

### 3.1.3 探究 DNS

回顾之前的讨论，注意到 IPv6 主机策略表，显然，在双栈主机中 DNS 给 IPv6 数据流扮演了重要的角色。为同样的主机域名分别存入 AAAA 记录和 A 记录，会使源设备在 IPv6 可用时均使用 IPv6 连接。事实上，对于本章所讨论的每个过渡技术，DNS 都是非常关键的，毕竟它对终端的不同域名标识之间提供了至关重要的联系，如应用层上的网站地址和目的 IP 地址（无论是 IPv4 或 IPv6 地址）。而且，由于 IPv6 需要编址，在应用层上，用户会发现，输入 IPv6 地址比较麻烦。终端用户若要访问双栈设备，需先进行 DNS 域名解析（可由管理员配置，分别将此节点的 IPv4 地址和 IPv6 地址分别关联一个 A 记录和一个 AAAA 记录）。记录中，域名所有者可能有相同或不同的关联到此设备的主机名，参见下面的例子：

dual- stack- host. ipamworldwide. com. 86400 IN A 10. 200. 0. 16
dual- stack- host. ipamworldwide. com. 86400 IN AAAA 2001 :db8 :2200 : : a

这个例子中，主机 dual- stack- host. ipamworldwide. com 对 IPv4 和 IPv6 均支持。双栈设备访问这台主机时，会先试图通过 IPv6 连接，若失败则转而试图进行 IPv4 连接。

为使 IP 地址转向主机，域名应该在 DNS 中使用恰当的 . arpa 域名进行配置。PTR 资源记录对应 IPv4 地址，列首写入恰当的 in- addr. arpa zone 文件；同理，IPv6 PTR 记录则在恰当的 ip6. arpa zone 文件中配置。而所谓“恰当的”zone 文件，则取决于组织的 DNS 管理策略。一些 DNS 集中化管理的组织会分别提供一个包含所有 IPv4 的 PTR 记录的 in- addr. arpa zone 文件，以及一个包含所有 IPv6

㊀ 实际上，在连接的初始化和测试中，IPv6 地址单独使用一个主机域名，对解决连接问题而言是一个好习惯。但是已被证实的是，AAAA 记录所有者将其转换为常规的主机名，而独立 IPv6 主机名可以被省去。

PTR 记录的 ip6. arpa zone 文件。而另一些组织则向子网管理者提供 DNS 权限，为每个子网都发放反向解析文件！虽然如此，大多数组织最后管理的 zone 文件数量都是适度的。不论有多少 zone 文件，都需要分别为每台主机配置资源记录。

16. 0. 200. 10. in-addr. arpa. 86400 IN PTR dual-stack-host.

ipamworldwide. com.

a. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 2. 2. 8. b.

d. 0. 1. 0. 0. 2. ip6. arpa. 86400 IN PTR dual-stack-host.

ipamworldwide. com.

要进行 DNS 解析、与目的主机（依照地址选择策略表返回的结果所对应的地址和协议）通信，双栈的节点必须支持接收 A 记录和 AAAA 记录。根据 DNS 查询和结果所使用的 IP 版本，RFC 3901 ［因特网当前最佳实践（Internet Best Current Practice 91）］建议每个递归式 DNS 要么支持单 IPv4，要么支持 IPv4/IPv6 双栈。这个 RFC 文档也建议，每个 DNS zone 应由至少一个 IPv4 可达的 DNS 授权服务器进行维护。这些建议阐述了为相当一段时间内大量的单 IPv4 使用提供的向后兼容的解决方案。因此，如果你计划在你的子网实施双栈，那么子网内必须包括 DNS。

### 3. 1. 4 探究 DHCP

实施双栈时，DHCP 的使用相对简单——各个栈分别使用对应版本的 DHCP。也就是说，要获得 IPv4 地址，就使用 DHCP（v4）；而要获得 IPv6 地址或前缀，就是用 DHCPv6。不过，两种 DHCP 形式都需要补充一些配置信息，如使用哪台 DNS 或 NTP 服务器等。这两种协议栈的服务器的信息结合如果处理得不适当，可能会使客户端不正常工作。例如，如果两栈的 DHCP 都提供了 DNS 地址，那么，IPv4 和 IPv6 的优先顺序，或者混合的优先顺序都可能导致地址不能正常传送。对这个情况的担忧一直存在，就像 RFC 4477（即本书参考文献［37］）中提到的一样。不过，现行的标准是，为 IPv4 使用 DHCP 服务器，另为 IPv6 使用 DHCPv6 服务器，当然实际上它们有可能运行在同一台普通的服务器上。

## 3. 2 隧道方法

IPv4/IPv6 共存技术的第二个主要类别是隧道技术，目前已经有支持在 IPv6 上的 IPv4 和在 IPv4 上的 IPv6 的多种隧道技术。这些技术总体来看可分为手工配置隧道或自动配置隧道。手工配置隧道是预先定义的，而自动配置隧道（也被

称为“软线隧道”）的建立和拆除是动态的。回顾一些隧道技术的基础后，下面将讨论这两种隧道类型。

一般情况下，在 IPv4 网络上的 IPv6 隧道技术需要将 IPv6 的报文加上 IPv4 的报头作为前缀。这样使得 IPv6 的数据报能够在 IPv4 路由设施上按照路线发送。IPv6 报文被简单地认为是 IPv4 数据报（见图 3-3）中的有效负载。隧道的入口节点，无论是路由器或主机，将进行封装。在 IPv4 报头的源地址处填入该节点的 IPv4 地址，然后目的地址是隧道另一个端点。IPv4 报头中的协议字段（protocol field）被设置为十进制的 41，用以标明这是一个 IPv6 报文的封装。在隧道的出口节点，IPv4 的报头将被剥离解封，该数据报将作为 IPv6 报文被路由至最终的 IPv6 目的地址。

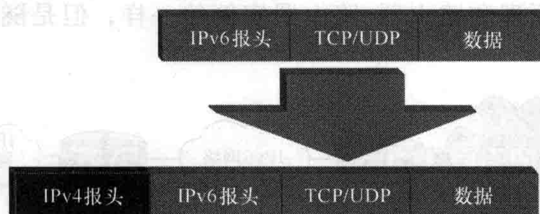



图 3-3　IPv6 上的 IPv4 隧道[28]

### 3.2.1　IPv4 网络上 IPv6 数据报的隧道方案

通过使用这种基本的隧道方法，设计了基于隧道的两个端点多种情况。可能最常见的配置是一条路由器到路由器的隧道，如图 3-4 所示，这是对于配置隧道的最常用的方法。




图 3-4　路由器到路由器的隧道[28]

图中，左侧起始的 IPv6 主机具有的 IPv6 地址为 W（为了简单起见）；去往 IPv6 地址为 Z 的远端主机的数据报[⊖]被发送到一个连接 W 所在子网的路由器上。

⊖ 图 3-4 中，报文大致标识为原始主机下的实线矩形，显示包的 IPv6 源地址 W 和目的地址 Z。在这个和随后的隧道图中，隧道头显示为虚线矩形。

该路由器的 IPv4 地址为 B，IPv6 地址为 X，接收 IPv6 报文。路由器被配置为将发送往 Z 所在的网络的数据报使用隧道传输，路由器将会用 IPv4 报头封装这些 IPv6 报文。该路由器使用其 IPv4 地址（B）作为 IPv4 源地址，然后用隧道终点的路由器的 IPv4 地址（C）作为目的地址，如图 3-4 所示的中央 IPv4 网络下的虚线矩形。使用隧道技术的数据报会像“常规”IPv4 数据报一样按照路由路径发送到隧道终端的目的路由器。该路由器解封装数据报，去除掉 IPv4 报头，然后将原始的 IPv6 数据报发送给原定的目的地址 Z。

另一种隧道技术方案是用能够同时支持 IPv4 和 IPv6 的一台 IPv6/IPv4 主机；该主机可以通过隧道技术将数据报发送给路由器；该路由器解封装数据报然后通过纯 IPv6 网络进行路由传输。这个流程和报头地址如图 3-5 所示。这种隧道技术的原理和路由器-路由器实例的一样，但是隧道的终点是不一样的。

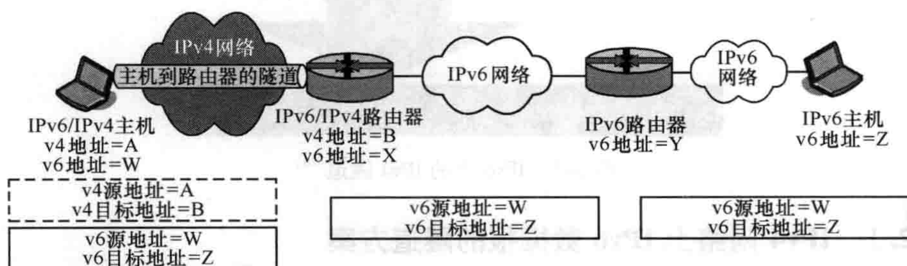



图 3-5　主机到路由器的隧道配置[28]

路由器-主机的隧道配置也很相似，如图 3-6 所示。图中左侧起始的 IPv6 主机发送 IPv6 数据报给本地路由器，本地路由器发送给距离目的地最近的路由器。该路由器被配置为利用隧道技术将 IPv6 数据报通过 IPv4 网络发送给图中所示的主机。

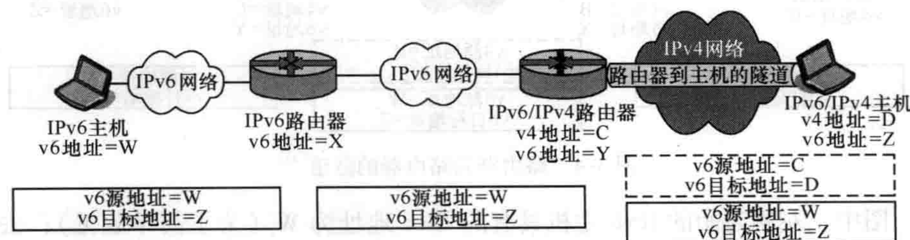



图 3-6　路由器到主机的隧道配置[28]

最后一种隧道技术的配置是跨度从头到尾，从主机到主机。如果路由基础设施还没有被升级到支持 IPv6，那么这种隧道技术的配置使得两台 IPv6/IPv4 主

机可以通过图 3-7 所示的隧道进行通信。

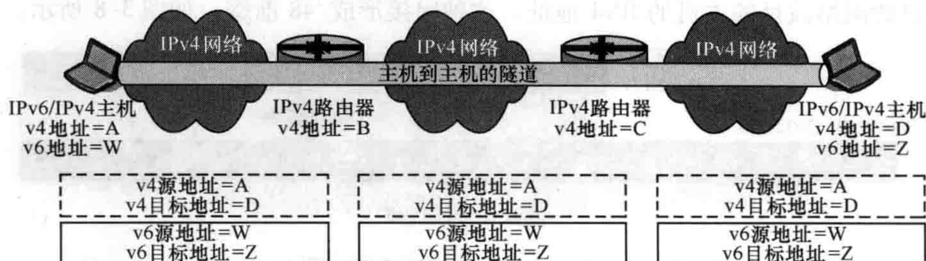



图 3-7 主机到主机的隧道配置[28]

### 3.2.2 隧道类型

正如之前所提到的，隧道可以是手工配置的也可以是自动配置的。手工配置的隧道，如 6in4 隧道，是管理员在通信前预先配置的。在上文描述的情境中，需要手动配隧道出入端点的隧道配置参数及其他参数，以决定何时使用隧道，也就是利用目的地址，来判别是否需要使用隧道。

自动配置隧道并不需要预先配置隧道，但还是要求进行配置启动隧道。隧道是基于包含了 IPv6 数据报（如源地址 IP 或目的地址 IP）在内的信息创建的。本节将介绍自动配置隧道技术：

• 6to4。自动配置隧道技术基于全局唯一的地址前缀和关联的全局（公共）IPv4 地址；6to4 可用于在基于 IPv4 的 MPLS 网络上互连企业网络中多个远程站点，虽然连接到因特网站点时，测量到的连接失败率为 10% ~25%[38]。

• 站内自动隧道寻址协议（Intra-Site Automatic Tunnel Addressing Protocol, ISATAP）。自动化主机-路由器，路由器-主机，或者主机-主机隧道技术，这基于一种特殊内嵌有 IPv4 地址的 IPv6 地址格式。

• 6over4。使用 IPv4 多播实现主机-主机自动配置的隧道技术。

• 隧道代理。当主机需要隧道时，由服务器作为隧道代理分配隧道网关资源的方式自动配置隧道。

• Teredo。通过在 IPv4 网络 NAT 防火墙的自动配置隧道，但测得的网络连接失败率约占连接尝试的 40% ~50%。

• 双栈过渡机制（Dual Stack Transition Mechanism，DSTM）。IPv4 数据报在 IPv6 网络上使用的自动配置隧道。

#### 3.2.2.1 6to4

6to4 是指一种“IPv6 over IPv4”的隧道技术。这种技术利用一种特殊的 IPv6 地址格式来标识 6to4 数据报。该地址格式包含了一个 6to4 的前缀，2002::/

16；其后为一个全球唯一 IPv4 地址，当使用主机/路由器-主机隧道技术时，则是目的网站或目的主机的 IPv4 地址。这种串接形成/48 前缀，如图 3-8 所示。

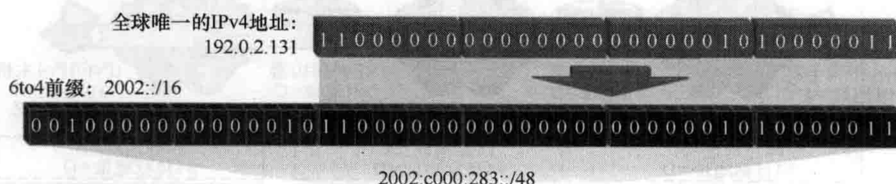



图 3-8 6to4 地址前缀的推导[39]

在这个例子中，唯一的 IPv4 地址 192. 0. 2. 131 表示终止 6to4 隧道的 6to4 路由器或主机的公共 IPv4 地址。48 位的 6to4 前缀作为全局路由前缀，一个子网 ID 可以被附加为接下来的 16 位，后跟一个完全定义的 IPv6 地址的接口 ID。使用 6to4 隧道技术的主机必须将其 IPv6 数据报的有效负载封装和解封装在 IPv4 数据报中（即追加 IPv4 报头），这样路由器可以在 IPv4 网络中对 IPv6 数据报进行路由寻址。该技术必须采用支持 6to4 隧道技术的路由器（6to4 路由器）；而且通过 6to4 隧道发送/接收数据的 IPv6 主机，必须使用 6to4 地址进行配置，该主机被认为是 6to4 主机。

来看一个例子：两个包含 IPv6 主机的网站想要通信，并通过与公共 IPv4 网络（如因特网）相连的 6to4 路由器互相连接。如图 3-9 所示，面向互联网的两台路由器的 IPv4 接口的 IPv4 地址分别为 192. 0. 2. 130 和 198. 51. 100. 1。将 IPv4 地址转化为 6to4 地址，分别得到了 2002：c000：282：：/48 和 2002：c633：6410：：/48。每个从各自网站来的/48 块都可以被通告在互联网上，可以在各自的组织内分配（分割）为/64 子网，对于每个子网都要求 IPv6 主机在 IPv4 互联网或其他 IPv4 私网上的连通性。为简单起见，左端的 6to4 主机子网 ID = 0 和接口 ID = 11。因此，这台主机的 6to4 地址为 2002：c000：282：1：：11。同样，在另一端（右端）的 6to4 主机所在子网 ID = 0 和接口 ID = 0，因而 6to4 地址为 2002：c633：6401：：F0。

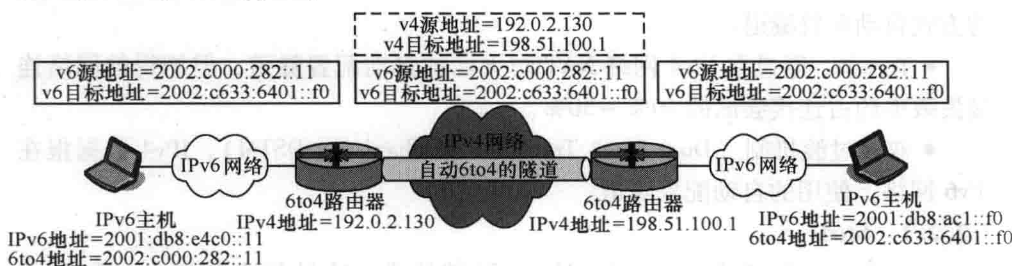



图 3-9 6to4 隧道的例子[39]

这些 6to4 地址对应的 AAAA 和 PTR 资源记录也应该被添加到相应域的 DNS 中。当使用隧道转发数据穿透因特网时，AAAA 和 PTR 所记录的目的地由每个管理相应 6to4 设备的组织维护，该决议可能需要向下遍历每个域的子树。该 AAAA 记录遵循正常“转发域”决议，但 PTR 记录不是那么简单。尽管 PTR 记录的域树形结构基于其对应的 IPv6 地址，但在 6to4 隧道中，它是由一个基于 IPv4 地址空间的机构“自我配置”的，而不是上游的 IPv6 地址管理组织分配，ip6. arpa 代理与上层的管理域无关。一个特别的注册项被设立来专门处理来自 2. 0. 0. 2. ip6. arpa 域［号码资源组织（Number Resource Organization，NRO）］的委托 。对应例子中 2002：c000：282：：/48 前缀的 ip6. arpa 域的管理员，将用 6to4. nro. net 连同相应的权威域名服务器注册 2. 8. 2. 0. 0. 0. 0. C. 2. 0. 0. 2. ip6. arpa 域。

回到数据报流，当图 3-9 所示左端的主机想要与右端的主机通信，会进行一次关于右端主机的 DNS 查找，在查找结果中将包含一条 6to4 地址解析（2002：c633：6401：：f0）。基于策略表里标签的匹配，发送信息的主机将利用其 6to4 地址（2002：c000：282：：11）作为源地址和目标主机 6to4 地址作为目的地址。当左侧的 6to4 路由器收到这个数据报，这个路由器将分别使用它的（源地址为 192. 0. 2. 130）和另一端 6to4 路由器的（目的地址为 198. 51. 100. 1）IPv4 地址作为 IPv4 报头封装数据报。接收到数据报的目的 6to4 路由器解封装去除 IPv4 报头，并将数据报路由到 2002：c633：6401：：/ 64 网络中的目标 6to4 主机。

6to4 可以为 IPv6 在 IPv4 网络的通信提供一个有效的机制。随着 IPv6 网络的逐步部署，6to4 中继路由器，即支持 6to4 的 IPv6 路由器，可以通过 IPv4 网络，从“纯”IPv6 网络主机到 IPv6 主机来传递数据报，实质上是充当 IPv6 网络和 6to4 主机或路由器之间的网关。这使得 6to4 主机可以与纯 IPv6 网络主机通信，反之亦然。

然而，对于这种寻址和隧道模式的应用，6to4 主机或路由器要求 6to4 中继路由器有能力将全球单播（本机）IPv6 地址转换成 6to4 隧道地址。这里有以下三种方法配置中继路由器：

1. 将 6to4 中继路由器作为目的 IPv6 网络地址的下一跳来配置路由器。

2. 利用普通的路由协议，使 6to4 中继路由器能够广播路由路线给 IPv6 网络。当广播的路由是指向迁移网络或者内部 IPv6 网时，可使用这种方式。如果说纯 IPv6 网络（见图 3-10）是“IPv6 网络”，那么下面的默认路由选项可能是更好的选择。

3. 配置一条通过 6to4 中继路由器访问 IPv6 网络的默认路由。这个方法可能应用于 IPv6 网络的连接只能通过组织内部的 IPv4 网络到达，或者这些组织内几乎没有纯 IPv6 网络。

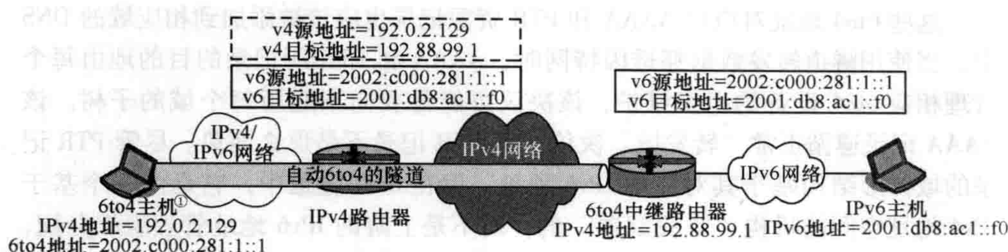



图 3-10 6to4 主机和纯 IPv6 主机通信[39]

① 原书误为 6to 主机。——译者注

如图 3-10 左端所示，得到一个在 IPv4/IPv6 网络上的 6to4 主机，这个主机的 6to4 地址是基于用户配置的公共 IPv4 地址（192. 0. 2. 129）配置的。该主机还必须配置路由或者默认路由去发送包含 6to4 任播地址（192. 88. 99. 1）或者包含 6to4 中继路由的 IPv6 地址的 IPv6 数据报。这台主机想要与使用 2001：db8：ac1：：f0 地址在纯 IPv6 网络上的主机进行通信，如图右边所示。当请求目的主机的 IP 地址时，DNS 返回的 AAAA 资源记录包含了该主机的 IPv6 地址。

因此，在左边的 6to4 主机生成的 IPv6 数据报是以其 6to4 地址为其源 IP 地址，以 IPv6 地址为其目的地址。然后主机用唯一的 IPv4 地址作为源地址，用对应的 6to4 中继器的 IPv4 地址作为目的地址来包装这个数据报，使之能够穿透 IPv4 网络。当 6to4 中继路由器的数据报到达时，中继路由器将数据报发送到 2001：db8：ac1：：/48 网络上。在相反方向，使用原来的源 6to4 地址作为目的 IPv6 地址，这样可以通知 6to4 中继路由器这个数据报需要 6to4 隧道到相应的目的主机。

#### 3.2.2.2 站内自动隧道寻址协议

ISATAP 是一个为主机到路由器、路由器到主机和主机到主机配置实现自动搭建跨越 IPv4 网络的 IPv6 隧道的实验性协议。ISATAP 的 IPv6 地址使用 IPv4 地址定义它的接口标识符。这个接口标识符由：：5efe：w. x. y. z 构成，其中 w. x. y. z 是一个点分十进制记法的 IPv4 地址。所以一个对应于 192. 0. 2. 131 的 ISATAP 接口标识符记为：：5efe：192. 0. 2. 131。使用 IPv4 的地址记号法可以明确地指示出 ISATAP 地址中所包含的 IPv4 地址，而不用把 IPv4 地址转换成十六进制的形式。ISATAP 接口 ID 可当作是一个普通的接口 ID，可以把它添加到它所支持的网络前缀后面去定义一个 IPv6 地址。例如，应用了上述 ISATAP 接口 ID 构成的 IPv6 本地链路地址是 fe80：：5efe：192. 0. 2. 131。

支持 ISATAP 的主机需要维护一个潜在路由器列表（Potential Router List，PRL），该列表包含了每一个可以广播 ISATAP 接口的路由器的 IPv4 地址和与之相关的地址生命周期计时器。通过 IPv4 网络上的路由请求，ISATAP 主机从本地路由器中请求到 ISATAP 支持信息。请求的目的端需要事先在主机上手动配置辨

别标识，或者是通过查找 DNS 中有主机名为“isatap”的路由器得到识别，又或者是使用 DHCP 供应商特定的能标识出 ISATAP 路由器 IPv4 地址的特殊选项进行鉴定辨别。DNS 技术要求管理员使用“isatap”这样的主机名来给 ISATAP 路由器创建资源记录。

一台 ISATAP 主机采用了图 3-11 所示的 IPv4 头部去封装 IPv6 数据报，其中使用的对应于被选定路由器的 IPv4 地址可从 PRL 中得到。

不管 IPv4 地址是静态定义的还是通过 DHCP 得到的，ISATAP 都能用已配置好的 IPv4 地址去自动配置它们的 ISATAP 接口 ID。如果配置了 IPv6，微软 Windows XP 系统和 Windows 2003 服务器就可以完成这样的自动配置。微软 Windows Vista 和Windows 7 系统及 Windows 2008 服务器默认支持 ISATAP 自动配置。ISATAP 接口 ID 可被添加到 64 位的全局网络前缀和子网标识符后面，由被请求的 ISATAP 路由器进行路由广播。

如图 3-11 所示，左端的主机已知目的主机的 IP 地址，此例中用的是 IPv6 地址。

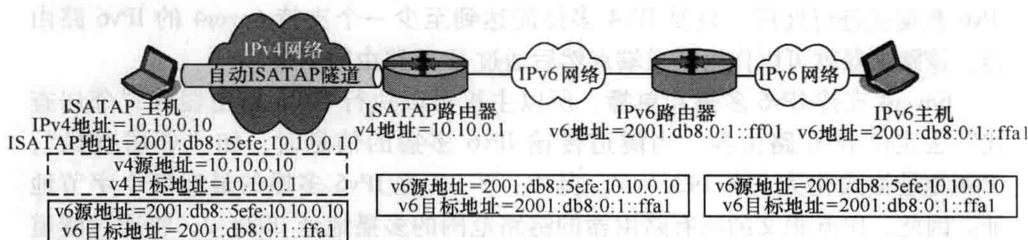



图 3-11　ISATAP 主机到路由的例子[39]

IPv6 数据报由主机构建，使用本主机的 ISATAP IPv6 地址作为源地址，目的 IPv6 主机的地址作为目的地址。这个数据报被封装在 IPv4 报头，从而形成一个自动隧道。隧道的源地址设置为 ISATAP 主机的 IPv4 地址，目的地址设置为 ISATAP 路由的 IPv4 地址，并且 IP 头部中的协议域设为十进制的 41 来指明这是一个被封装的 IPv6 数据报。ISATAP 路由器不需要和主机处在相同的物理网络，而且这个隧道可以跨越主机和 ISATAP 路由器之间 IPv4 网络（0 跳或多跳）。ISATAP 路由器除去 IPv4 头部后，对剩下的 IPv6 数据报进行正常的 IPv6 路由选择以送往目的主机。

目的主机会使用源主机的 ISATAP 地址向源主机发出响应。因为 ISATAP 地址包含了一个全局唯一的网络前缀/子网 ID，目的数据报会被发往提供隧道服务的 ISATAP 路由器。通过处理接口 ID，本地的 ISATAP 路由器可以提取出目的主机的 IPv4 地址，并且向源主机发出用 IPv4 头部封装的 IPv6 数据报。图 3-11 中，从右到左，ISATAP 路由器可以构建通向目的主机的 ISATAP 隧道。

主机到主机的 ISATAP 隧道和图 3-7 所示的相似，可由 IPv4 网络上的 ISATAP 主机发起来构建，条件是本地链路（同一个子网）或者全局网络的前缀可以添加到主机的 ISATAP 接口 ID 的前面。如图 3-7 所示，ISATAP 地址是分别由 IPv4 地址 A 和 D 生成的 IPv6 地址 W 和 Z 。

**3.2.2.3 6over4**

6over4 是一个利用 IPv4 多播的自动隧道技术。对于 6over4，IPv4 多播是必需的，它被认为是一个虚拟链路层或虚拟以太网。从虚拟链路层的角度来看，IPv6 的地址是由本地链路范围前缀标识的（fe80::/10)。一个主机的 IPv4 地址包括其 IPv6 地址中 6over4 接口 ID 的部分。例如，一个 IPv4 地址为 192.0.2.85 的 6over4 主机的 IPv6 接口 ID 为::c000:255，因此，6over4 的地址为 fe80::c000:255。6over4 隧道可以是包括主机-主机、主机-路由器、路由器-主机等类别。在各种情况中，主机和路由器都需要分别配置支持 6over4。IPv6 数据报在隧道中使用 IPv4 报头相应的 IPv4 多播地址。该多播组的所有成员都会收到隧道的数据报，就像是在虚拟链路层中一样，真正的接收者会去除掉 IPv4 报头再对 IPv6 数据报进行处理。只要 IPv4 多播能达到至少一个支持 6over4 的 IPv6 路由器，该路由器就可以作为隧道端点然后通过 IPv6 路由数据报。

6over4 支持 IPv6 多播和单播，所以主机可以执行 IPV6 路由器和“邻居查找”去定位 IPv6 路由器。当隧道传输 IPv6 多播的信息时，如“邻居查找”，IPv4 地址格式化为 239.192.Y.Z。其中，Y、Z 是 IPv6 多播的最后两个字节地址。因此，IPv6 报文的所有路由器的链路范围的多播地址 ff02::2，将通过隧道送至 IPv4 目的地 239.192.0.2。6over4 主机可使用网络组会员协议（Internet Group Membership Protocol，IGMP）来告知 IPv4 路由器关于多播组成员的信息，所以路由器能将多播数据报转发给它们。

**3.2.2.4 隧道代理**

隧道代理是另一种 IPv4 网络上实现隧道自动管理的技术。隧道代理负责管理两类对象：一类是来自双栈客户端的隧道请求，一类是连接着目的 IPv6 网络的隧道服务器。想要连接至 IPv6 网络的双栈的客户端被有选择性地导向一个隧道管理的门户网站去进行身份认证，认证通过后便可以使用隧道代理服务。隧道代理也可能会自己管理认证和授权。客户端需要提供自己的 IPv4 地址、自己想要的 FQDN、IPv6 地址的请求数、自己的类型（主机还是路由器）等信息。

一旦验证通过，隧道代理就会执行以下工作去创建隧道：

1. 分配和配置一个隧道服务器，并告知其客户端的信息。
2. 根据请求的地址数量和客户端类型分配 IPv6 地址或前缀给客户端。
3. 在 DNS 上注册客户端的 FQDN。
4. 提供分配的隧道服务器和相关的隧道及 IPv6 参数（包括地址/前缀，

DNS 名称）等信息给客户端。

图 3-12 所示的隧道代理的交互，说明了客户端-隧道代理的交互过程。RFC 5572 规范化了隧道建立协议（Tunnel Setup Protocol，TSP）以促进隧道建立的过程和组件之间的交互。

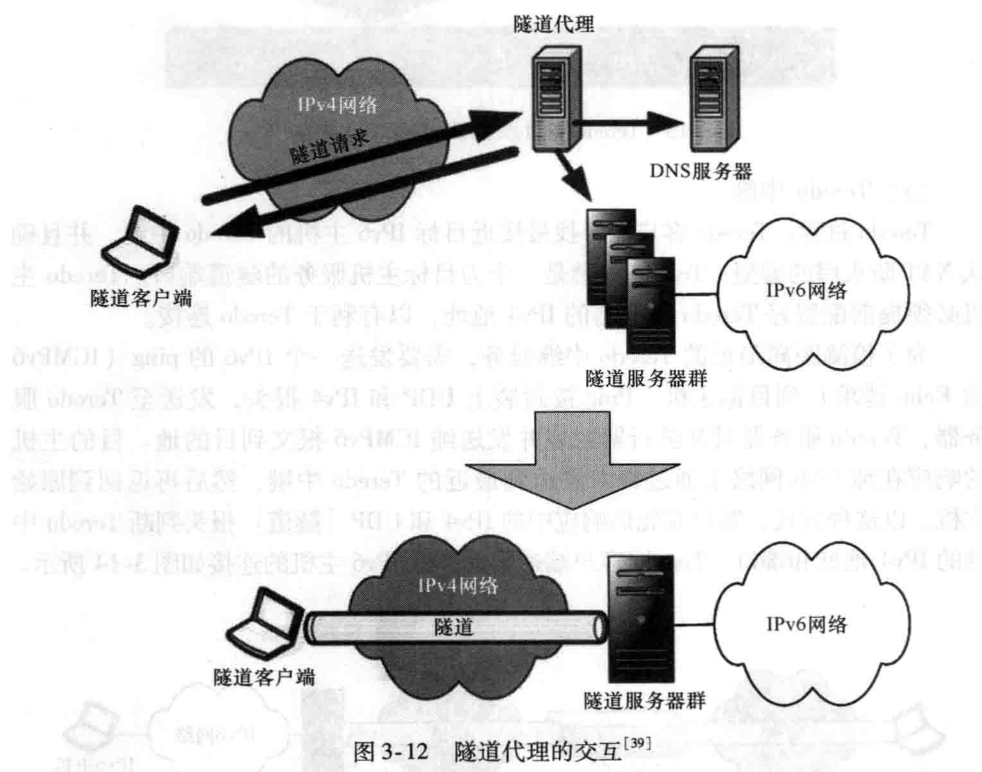



图 3-12 隧道代理的交互[39]

*Teredo* 实现穿透采用 NAT 防火墙的隧道技术是很难的，除非是经过各种复杂的设计。Teredo 为通过 IPv4 的 IPv6 数据报提供基于 UDP 的 NAT 穿透，可实现主机到主机的自动隧道功能。Teredo 通过使用附加的 UDP 报头，使得 NAT/防火墙之间的传送更加方便。很多 NAT/防火墙设备是不允许头字段为 41（前面提到的 IPv6 的数据报的隧道设置）的 IPv4 数据报传输的。这个附加的 UDP 报头掩盖了这些，使得被封装的 IPv6 可以穿透 NAT 防火墙设备，这种 UDP 端口的转换，大部分设备都可以支持（见图 3-13）。

Teredo 在 RFC 4380 中被定义，它被当作 IPv6 连接最后的转换技术。因为随着 6to4 及支持 IPv6 的防火墙路由器的部署，它的应用将越来越少。如图 3-14 所示，Teredo 需要以下部件：

（1）Teredo 客户端

（2）Teredo 服务器

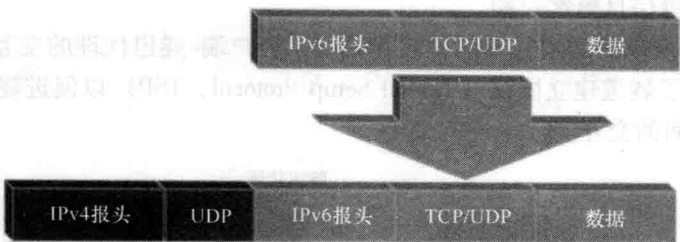



图 3-13 Teredo 隧道添加 UDP 和 IPv4 报头[39]

（3）Teredo 中继

Teredo 过程：Teredo 客户端寻找最接近目标 IPv6 主机的 Teredo 中继，并且确认 NAT 防火墙的类型。Teredo 中继是一个为目标主机服务的隧道端口。Teredo 主机必须提前配置好 Teredo 服务器的 IPv4 地址，以有利于 Teredo 连接。

为了检测距离最近的 Teredo 中继服务，需要发送一个 IPv6 的 ping（ICMPv6 的 Echo 请求）到目标主机。Ping 被封装上 UDP 和 IPv4 报头，发送至 Teredo 服务器，Teredo 服务器对其进行解封装并发送纯 ICMPv6 报文到目的地。目的主机的响应在纯 IPv6 网络上通过路由发送到最近的 Teredo 中继，然后再返回到原始主机。以这种方式，客户端凭借响应中的 IPv4 和 UDP［隧道］报头判断 Teredo 中继的 IPv4 地址和端口。Teredo 客户端通信到本地 IPv6 主机的连接如图 3-14 所示。

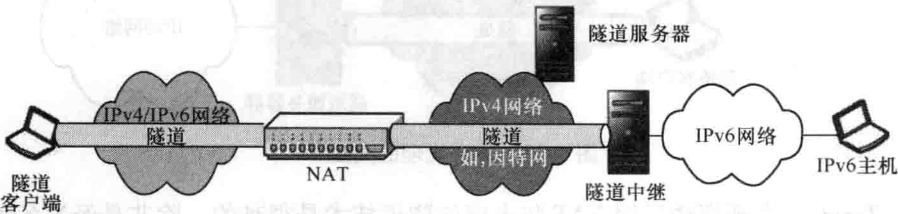



图 3-14 Teredo 客户端到 IPv6 主机的连接[39]

*NAT类型* 最初，Teredo 旨在解决穿透防火墙和对某些端口设限的 NAT。RFC 6081 定义了 Teredo 扩展，极大地扩展了支持穿透的 NAT 类型。表 3-2 列出了在 RFC 6081 中 Teredo 扩展支持的 NAT 类型[42]：

**表 3-2 RFC 6081 中 Teredo 扩展支持的 NAT 类型**

| | 目的地址的 NAT | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 圆锥形 | 地址受限 | 端口受限 | UPnP 端口受限 | UPnP 端口对称 | 端口保持端口对称 | 连续端口对称 | 端口对称 | 地址对称 |
| 圆锥形 | Yes | Yes | Yes | Yes | SNS | SNS | SNS | SNS | SNS |
| 地址受限 | Yes | Yes | Yes | Yes | SNS | SNS | SNS | SNS | No |

（续）

| | 目的地址的NAT | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 圆锥形 | 地址受限 | 端口受限 | UPnP 端口受限 | UPnP 端口对称 | 端口保持端口对称 | 连续端口对称 | 端口对称 | 地址对称 |
| 端口受限 | Yes | Yes | Yes | Yes | No | SNS + PP | SNS + PP | No | No |
| UPnP 端口受限 | Yes | Yes | Yes | Yes | SNS + UPnP | No | No | No | No |
| UPnP 端口对称 | SNS | SNS | No | SNS + UPnP | SNS + UPnP | No | No | No | No |
| 端口保持 | SNS | SNS | SNS + PP | No | No | SNS + PP | SNS + PP | No | No |
| 端口对称 | | | | | | | | | |
| 连续端口对称 | SNS | SNS | SNS + SS | No | No | No | No | No | No |
| 端口对称 | SNS | SNS | No | No | No | No | No | No | No |
| 地址对称 | SNS | No | No | No | No | No | No | No | No |

其中：

- Yes 为被初始的 Teredo 规范支持（RFC 4380）。
- SNS 为被对称 NAT 扩展支持。
- SNS + UPn 为被对称 NAT 扩展支持和通用即插即用（UPnP）的对称 NAT 扩展支持。
- SNS + PP 为被对称 NAT 扩展支持和端口保持对称 NAT 扩展支持。
- SNS + SS 为被对称 NAT 扩展支持和连续端口对称 NAT 扩展支持。
- No 为不支持。

要使用 Teredo 通信时，NAT 必须“初始化”以正确映射那些在 NAT 内的源地址和目的地址。为了完成 NAT 上关于内部主机与目标主机之间通信的映射，Teredo 客户端会发送一个气泡数据报到目的主机。气泡数据报是没有有效负载的 IPv6 报头，本身封装在 Teredo 的隧道 IPv4/UDP 头中。它使 NAT 可以完成内部地址和外部 IP 地址映射，以及内部端口号和外部端口号的映射。

一般来说，气泡数据报是直接从源 Teredo 客户端发送到目的主机。但是，如果目标主机位于防火墙后面，气泡数据报可能会被丢弃，因为这是不请自来的外部数据报。在这种情况下，Teredo 客户端会超时，然后通过 Teredo 服务器来发送气泡数据报，这个数据报可以通过 Teredo 格式的 IPv6 目的地址来识别，这个 IPv6 地址包含了目标主机的 Teredo IPv4 地址。然后 Teredo 服务器将 Teredo 数据报通过隧道转发到目标主机。主机有自己的 IPv4 地址，这个地址也被编码

在 Teredo IPv6 地址中。

如果目标主机也是一个 Teredo 客户端，它也将能接收到数据报。Teredo 客户端通过之前发送到 Teredo 服务器的 ping 请求完成 Teredo 客户端的初始化配置。之后，目标主机会直接向源主机响应，完成 NAT 映射（在目标主机和源主机双边都完成映射）。如图 3-15 所示，有两个 Teredo 客户端通过一个共同的 Teredo 中继通信。

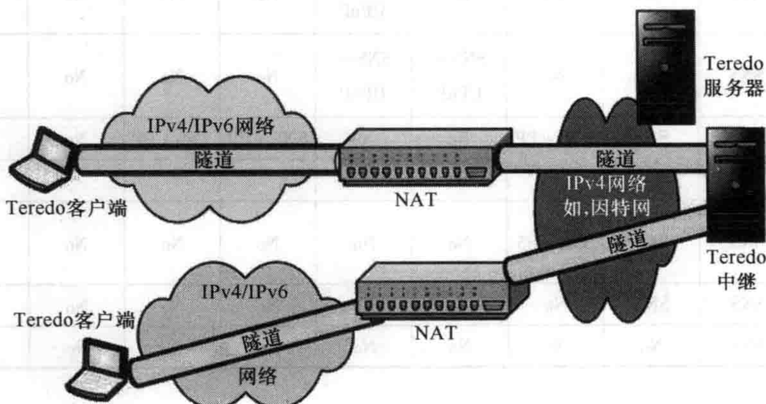



图 3-15　Teredo 客户端通过 IPv4 因特网连接[39]

正如已经看到的，Teredo 的 IPv6 地址由客户端及其 Teredo 服务器的 IPv4 地址来生成。图 3-16 给出了 Teredo IPv6 地址格式。

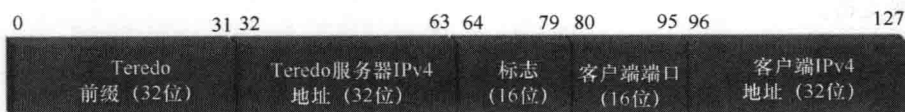



图 3-16　Teredo IPv6 地址格式[39]

Teredo 的前缀是一个预定义的 IPv6 前缀：2001::／32。Teredo 服务器 IPv4 地址构成下一个 32 位。虽然 RFC 5991（即本书参考文献［44］）将原来的 RFC 4380 定义的标志字段重新定义了，引进一个随机字符串来免去 cone 位的设置，在某些情况下它仍是被解释的。Teredo 标志字段格式如图 3-17 所示。




图 3-17　Teredo 标志字段格式

- C = cone 位
- z = 保留（设置为 0）

- Random1 为随机位 1
- U = 全局/局部位（设置为 0）
- G = 个人/全局位（设置为 0）
- Random2 为随机位 2

如果 U 和 G 位被置零，可以表明一个本地管理的单播地址和随机位，以阻止 Teredo 的 IPv6 地址扫描，即使对应的 IPv4 地址是已知的（对于一个给定的 IPv4 地址映射，12 个随机位可以提供 4096 种组合的 Teredo 地址）。客户端端口与客户端的 IPv4 地址字段通过反转每个位的值来表示这些字段对应的含义。

### 3.2.3 IPv6 网络上的 IPv4 数据报的隧道方案

启用 IPv6 后，一些 IPv6 用户仍旧需要与 IPv4 应用或 IPv4 网络上的主机进行通信，如因特网。IPv4 Over IPv6 的隧道技术为这种通信提供了一种实现方法。

#### 3.2.3.1 双栈过渡机制

DSTM（双栈过渡机制）是一种可以在 IPv6 网络上将 IPv4 数据报传输到目的地 IPv4 主机的隧道代理方法，如图 3-18 所示。IPv6 的主机如果需要与 IPv4 主机通信时，需要配置双协议栈及 DSTM 客户端。在对一个目的主机的主机名使用 DNS 获得其 IPv4 时，客户端会开始启动 DSTM 进程。这与隧道代理技术十

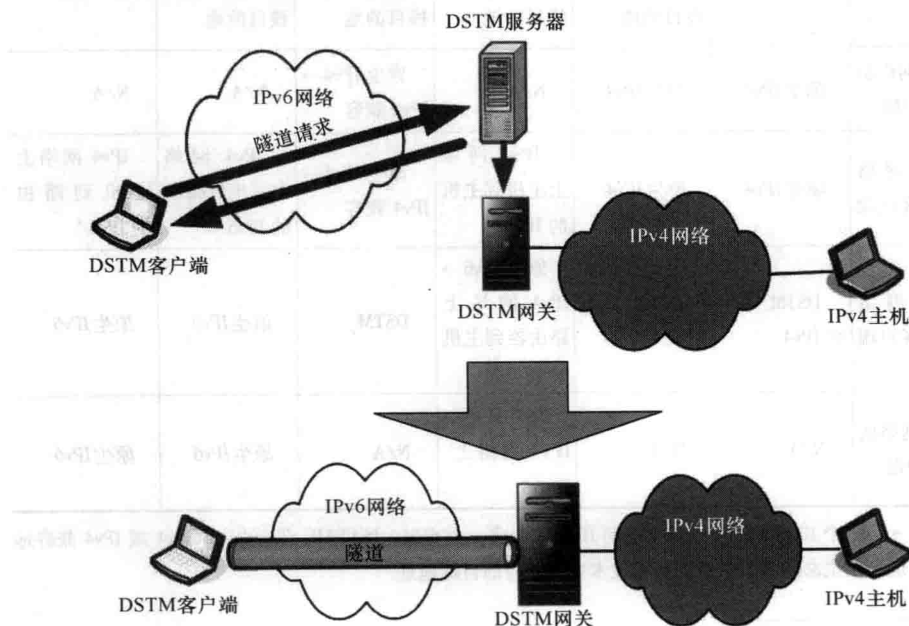



图 3-18 DSTM 隧道设置[39]

分类似。当整个进程开始，DSTM 客户端联系 DSTM 服务器通过 DHCPv6 协议[⊖]以获得一个 IPv4 地址，以及 DSTM 网关的 IPv6 地址。IPv4 地址被当作源地址在数据报里被传送。这个数据报利用 DSTM 客户端源 IPv6 地址和 DSTM 网关的 IPv6 地址作为目的地址进行封装。在 IPv6 报头中的“下一头部”字段可以用来表示这是一个使用“4over6”隧道技术封装的 IPv4 分组。

一种 DSTM 的变体可以支持使用 VPN 从本地网络外部进行访问的 DSTM 客户端，如在家工作的职员。在这种情况下，假设 DSTM 客户端获得一个 IPv6 地址而并非 IPv4 地址，它可以联络 DSTM 服务器以获得一个 IPv4 地址，这种访问需要认证以建立 DSTM 客户端与 DSTM 网关之间的一个 VPN。

### 3.2.4 隧道技术总结

表3-3 给出的隧道技术总结，说明了各种隧道技术的适用对象，分别基于源主机能力/网络类型、目标地址解析方式及网络类型等。

**表3-3 隧道技术总结**

| 目标网络 / 源网络 | IPv4 网络的 IPv4 目的地 | IPv4 网络上解析成 IPv4 地址的双协议栈目的地 | IPv4 网络上解析成 IPv6 地址的双协议栈目的地 | IPv6 网络上解析成 IPv4 地址的双协议栈目的地 | IPv6 网络上解析成 IPv6 地址的双协议栈目的地 | IPv6 网络的 IPv6 目的地 |
|---|---|---|---|---|---|---|
| IPv4 网络的 IPv4 客户端 | *原生 IPv4* | *原生 IPv4* | N/A | 原生 IPv4→IPv4 兼容 | N/A | N/A |
| IPv4 网络的双栈客户端 | *原生 IPv4* | *原生 IPv4* | IPv4 网络上主机到主机的 IPv6* | 原生 IPv4→IPv4 兼容 | IPv4 网络上主机到路由的 IPv6* | IPv4 网络上主机到路由的 IPv6* |
| IPv6 网络的双栈客户端 | DSTM→原生 IPv4 | DSTM→原生 IPv4 | 原生 IPv6→IPv4 网络上路由器到主机的 IPv6* | DSTM | *原生 IPv6* | *原生 IPv6* |
| IPv6 网络的 IPv6 客户端 | N/A | N/A | 原生 IPv6→IPv4 网络上 IPv6* | N/A | *原生 IPv6* | *原生 IPv6* |

注：* 一个 IPv6 地址可以是原生的 IPv6 地址或一个 6to4、ISATAP、Teredo、6over4 或 IPv4 兼容地址，主机必须基于它所支持的技术选择相应的目标地址。

⊖ 虽然 DSTM RFC 草稿（即本书参考文献［45］）表示 DHCPv6 是获得一个 IPv4 地址的更好的方法，但是 DHCPv6 目前并不定义为主动或通过一个选项设置分配 IPv4 地址。

表格中斜体字部分表示端对端设备的 IP 版本。任何使用不同协议的中间网络，必须或通过一个路由器到路由器的隧道或在边界处使用转换技术进行转换，这一点将在之后的章节介绍。

表中其余部分表示用到隧道技术的各种场景。符号“→”代表一个转换点或者能将相应本地网络协议转换为隧道协议（反之亦然）的隧道端点。

表中“N/A”部分表示无法通过隧道建立有效连接，然而，转换技术可能用来弥补这些缺陷。这些将在以后讨论。

## 3.3 翻译策略

从 IPv4 到 IPv6 的翻译（反之亦然）是在协议栈的特定层中执行，典型层有网络层、传输层或者应用层。与不修改传输的数据报而仅是添加一到两个报头的隧道技术不同，翻译机制做的是修改。也就是说，将 IP 报文在 IPv4 和 IPv6 之间转化。当使用 IPv6-only 技术的节点与使用 IPv4-only 技术的节点需要交换数据的时候，翻译策略通常是不错的选择。也就是说，表 3-3 中“N/A”部分需要用到翻译策略。在双协议栈的情况下，原生或者隧道机制是更好的选择。

在 IPv6 规范形成阶段，早期发展的 IPv4/IPv6 翻译方法已经被证明是不一致的，并且在大多数场合是不安全的。基于前车之鉴，一系列新的 RFC 的发布明确了 IPv4/IPv6 的翻译方法、寻址与一致性的策略。RFC 6144（即本书参考文献[46]）定义了 IPv4/IPv6 翻译的框架，并且明确了这种翻译策略将应用于何种网络互连的环境。可行的翻译方法，见表 3-4。其中，总结的各种情况所对应使用的翻译策略都是有指导意义的。每种情况都模拟了不同的源网络是如何连接到目标网络的，源网络和目标网络可能是私有网络上的，或者是使用某种协议的互联网络中的一台主机。

表 3-4　可行的翻译方法[46]

| 情景 | 源 网 络 | 目 的 网 络 | 适 用 范 畴 |
|---|---|---|---|
| 1 | IPv6 网络 | IPv4 因特网 | 带 DNS64 的无状态翻译可行 |
| 2 | IPv4 因特网 | IPv6 网络 | 带特殊网络前缀的无状态翻译可行 |
| 3 | IPv6 因特网 | IPv4 网络 | 带特殊网络前缀和 DNS 中有可翻译的 IPv4 地址的有状态翻译可行 |
| 4 | IPv4 网络 | IPv6 因特网 | 翻译不可行 |
| 5 | IPv6 网络 | IPv4 网络 | 类似情景 1，可行 |
| 6 | IPv4 网络 | IPv6 网络 | 类似情景 2，可行 |
| 7 | IPv6 因特网 | IPv4 因特网 | 翻译不可行 |
| 8 | IPv4 因特网 | IPv6 因特网 | 翻译不可行 |

情景 4 和 8 在整个因特网范围内都无法将一个 IPv6 地址唯一地翻译为一个 IPv4 地址。情景 3 是可行的，有能力将一个 IPv4 网络的地址压缩成一个 IPv6 地址的前缀。情形 7 在整个因特网地址空间上不具有这种压缩能力。

## 3.3.1 IP/ICMP 翻译

既然已经介绍了 IPv4/IPv6 翻译可行的情况，现在就来探讨翻译的技巧吧。将 IPv4 和 IPv6 数据报相互转化的算法是在 RFC 6145 中阐述的 IP/ICMP 转换算法，这个算法实现了在主机或网关上将发送出去的 IPv6 数据报的报头转换为 IPv4 报头，将接收到的 IPv4 报头转换为 IPv6 的，反之亦然。可以认为在主机上实现的翻译算法（如“主机泵”，将在后文进行描述）与网关起相同作用，为了简化讨论，这里主要考虑网关如何进行翻译。翻译的过程涉及地址转换、数据报分割、ICMP 映射还有 IP 报头字段的转换。

### 3.3.1.1 地址翻译

地址翻译在 RFC 6052（即本书参考文献［48］）中被定义，并且可应用于任何需要将 IPv4 地址与 IPv6 地址相互转换的实体，包括网关转换和 DNS64 服务。语义上，一个转换 IPv4 的 IPv6 地址（IPv4-converted IPv6 address）是用一个 IPv6 地址代表一个 IPv4 节点，而 IPv4 可翻译的 IPv6 地址（IPv4-translatable IPv6 address）则是为了进行无状态地址翻译将一个 IPv6 地址分配给一个 IPv6 节点。其前缀就是翻译过程的输出，而后缀则是代表了一个预先映射好的内嵌 IPv4 地址的 IPv6 地址。

两者都有着一样的格式，由一个级联 32 位 IPv4 地址的 IPv6 前缀组成，并且大多数情况下跟有后缀。唯一需要注意的是，为了与 IPv6 寻址兼容，64 ~ 71 位（IID 的第 1 个八位组）被设为 0，并且指定了第 70 位作为“通用/本地”（u）标识位，第 71 位作为“个人/群组”（g）标志位，这些位置为 0 则表示本地执行的单播地址。根据不同 IPv6 前缀的长度，可以将 IPv4 地址与“U”字节（Byte）插入到图 3-19 所示的位置上。




图 3-19 IPv4 可翻译和转换 IPv4 的 IPv6 地址格式

IPv6 前缀的长度只能为图 3-19 所示的取值：32、40、48、56、64 和 96 位。96 位的前缀已经被分配给 64：ff9b：：/96，因为它只能代表唯一的公共 IPv4 地址并且通常只应用于组织的翻译服务。使用 96 位的前缀必须保证 U 字节全为 0。例如，可以在一个 64 位前缀的后面添加 32 个 0 从而获得兼容的 96 位前缀。不过，通常组织会从它们指定的地址空间中分配一个前缀专门用于表示可翻译的 IPv4 地址，前缀比总共分配到的地址多 8 位（假如在 256 位的地址空间分配到 48 位，则使用 56 位的前缀）。如果这一地址前缀还未在网络上进行通告的话，要将这地址前缀进行通告。翻译网关需要进行配置以便识别可翻译的 IPv4 地址前缀，并且将 IPv6 报文中内嵌的 IPv4 地址作为目的地址转换为 IPv4 报文。

例如，考虑一个 IPv4 地址为 198.51.100.49 的主机，它可以通过一个配置好的翻译网关，使用前缀 2001：db8：3a01：4f00：：/56 变成可到达的 IPv6 地址。将 IPv4 地址转换为十六进制得到 c633：6431，在保留 U 字节 0 位的前提下添加到前缀，可以得出这个可翻译的 IPv6 地址 2001：db8：3a01：4fc6：33：6431：：。如图 3-20 所示，这是左侧主机的 IPv6 地址。它在 DNS 中的可达性可以是 195.51.100.49（A 条目），也可以是 2001：db8：3a01：4fc6：33：6431：：（AAAA 条目）。同样的，右侧拥有 IPv4 地址 192.0.2.188 的主机也可以用 IPv4 转换的 IPv6 地址 2001：db8：3a01：4fc0：0：2bc：：表示。通过分析 IPv4 主机的信息，左边的主机发送了指定的报文到 2001：db8：3a01：4fc0：0：2bc：：。指定到 2001：db8：3a01：4f00：：/56 的报文被路由到 NAT64 网关，就是它起到了本章描述的 IP/ICMP 转换的功能，以及 IPv6 地址与 IPv4 地址的相互映射。如果右侧的主机是双协议栈并且可通过 IPv6 直达，翻译功能就会被忽略，该功能只有在没有原生协议存在的时候才会应用。

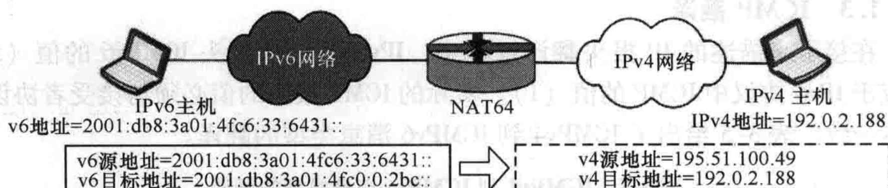



图 3-20　IP/ICMP 翻译例子

#### 3.3.1.2　数据报分段

数据报分段可以将一个大数据报分成两个甚至更多的小数据报，从而使其通过最大传输单元（Maximum Transmission Unit，MTU）比自身长度小的中间网络。在 IPv4 中，需要的话，路由器可以起到数据报分段的功能；而在 IPv6 中，数据报分段由节点单独完成，而不是路由器。IPv6 主机在进行传输前，为相应大小的数据报寻找合适的最小 MTU 路径。主机给目的地传输数据报的时候，首

先假设线路的 MTU 与本地连接的 MTU 相等；如果中间某一跳的 MTU 比数据报小的话，一条关于"数据报太大"的 ICMPv6 错误信息就会发送回主机，表明这是一条无效连接。主机可能会根据相应的 MTU 大小调整数据报大小；如果在传输过程中遇到更小的 MTU，这个步骤将会重复执行。IPv4 节点也可以通过发送所需的 MTU 数据报到目的地完成 MTU 路径发现，不过因为在 IPv4 中路由器可能会自动将大的数据报分段，主机可以在主 IPv4 报头设置"不分段"（Don't Fragment DF）位来禁用数据报分段。IPv6 的情况也类似，在传输过程中，如果数据报大小超过了 MTU，检测到该情况的路由器将会给主机发送一条"需要分段"的 ICMP 错误信息。

当翻译器接收到设置了 DF 位的 IPv4 数据报并且下一个 MTU（或者 IPv6 的下一跳收到"数据报太大"的 ICMP 信息）小于 IPv4 数据报大小 +20（考虑到 IPv6 报头增长的大小），一条"需要分段"的 ICMP 信息将会被发送到源 IPv4 地址。如果 DF 位没有被设置，但是数据报大小超过了下一个或下一跳的 MTU，翻译器就会将数据报分段。如果数据报小于 MTU，翻译器也可能根据配置添加一个分段报头，仅为了表明分段是允许的。如果 DF 位被设置了，除非 MTU 对于要发送的数据报足够大，否则永远不要添加分段报头。

当翻译器将接收到的 IPv6 数据报转换为 IPv4 时，它默认会设置 DF 位。如果接下来接收到"需要分段"的 ICMP 信息作为回复，它就会进而翻译"数据报太大"的 ICMPv6 信息并且把它发送回原始的 IPv6 主机。原始主机不需要使用数据报小于 1280 字节的最小 IPv6 MTU，但是它会转发带有分段报头的数据报，翻译器将会给每个传输到 IPv4 目的地的报头设置标识值。在这种情况下，不设置 DF 位表明下一个 IPv4 分段是允许的。

#### 3.3.1.3　ICMP 翻译

在接下来描述的 IP 报头翻译过程中，IPv6 下一个报头 ICMPv6 的值（58）对应于 IPv4 协议中 ICMP 的值（1）。实际的 ICMP 报头的值必须与接受者协议的版本一致。表 3-5 给出了 ICMPv4 到 ICMPv6 消息类型的翻译。

**表 3-5　ICMPv4 到 ICMPv6 消息类型的翻译**

| ICMPv4 消息类型 | 翻译的 ICMPv6 消息类型 |
| --- | --- |
| 回显（8）和回复（0） | 回应请求（128）和回应答复（129） |
| ICMP 路由通告/请求（9，10） | ICMPv6 已废弃 |
| 时间戳和时间戳回应（13，14） | ICMPv6 已废弃 |
| 信息请求/回复（15，16） | ICMPv6 已废弃 |
| 地址掩码请求/回复（17，18） | ICMPv6 已废弃 |
| ICMP 消息 | 抛弃 |

（续）

| ICMPv4 消息类型 | | 翻译的 ICMPv6 消息类型 | |
|---|---|---|---|
| 目的地不可达 | Code 0，1（网络，主机不可达） | | 目的地不可达（1），Code = 0（没有到达目的地的路由） |
| | Code 2（协议不可达） | | ICMPv6 参数问题（4），Code = 1（无法识别下个报头类型） |
| | Code 3（端口不可达） | | 目的地不可达（1），Code = 4（端口不可达） |
| | Code 4（需要分段数据报，设置 DF 位） | | 数据报太大（2），Code = 0，调整最大传输单元 |
| | Code 5（源路由失败） | | 目的地不可达（1）Code = 0（没有到达目的地的路由） |
| | Code 6，7，8（未知目的网络，未知目的主机，源主机被隔离） | | 目的地不可达（1）Code = 0（没有到达目的地的路由） |
| | Code 9，10（与目的主机/网络的通信被禁止） | | 目的地不可达（1），Code = 1（与目的地通信被禁止） |
| | Code 11，12（由于服务类型导致目的网络/主机不可达） | | 目的地不可达（1）Code = 0（没有到达目的地的路由） |
| | Code 13（通信被禁止） | | 目的地不可达（1），Code = 1（与目的地通信被禁止） |
| | Code 14（主机越权） | | 抛弃 |
| | Code 15（优先中断） | | 目的地不可达（1），Code = 1（与目的地通信被禁止） |
| 报源抑制（4） | | ICMPv6 已废弃 | |
| 重定向（5） | | 抛弃 | |
| 选择主机地址（6） | | 抛弃 | |
| 超时（11） | | 超时（3），有相同代码值 | |
| 参数问题 (12) | Code 0（指针显示错误） | 参数问题 (4) | Code = 0（遇到错误头字段）并且更新指针值或抛弃（按表 3-6） |
| | Code 1（缺少必要的选项） | | 抛弃 |
| | Code 2（错误长度） | | Code = 0（遇到错误头字段）并且更新指针值或抛弃（按表 3-6） |
| 其他代码值或未知的 ICMPv4 类型 | | 抛弃 | |

对于参数问题的错误信息，报头指针值的翻译见表 3-6。

**表3-6　ICMPv4到ICMPv6报头指针值的翻译**

| ICMPv4 指针值 | 翻译的ICMPv6指针值 |
| --- | --- |
| 版本/IP报头长度（0） | 版本/流量类型（0） |
| 服务类型（1） | 流量类型/流标签（1） |
| 总长度（2，3） | 负载长度（4） |
| 标志位（4，5），标识/分割位移（6，7），报头校验和（10，11） | 抛弃 |
| 生存时间（8） | 跳数限制（7） |
| 协议（9） | 下一个报头（6） |
| 源地址（12～15） | 源地址（8） |
| 目的地址（16～19） | 目的地址（24） |

ICMPv6校验和的计算必须在进行完整的翻译过程之后。ICMP错误的有效负载可能会导致翻译问题，特别是当其中包含了另一种协议的IP地址。如果附加的IP报头也是返回的错误信息的一部分，翻译网关通常会企图翻译这个附加的IP报文头。

表3-7总结了ICMPv6到ICMPv4消息类型的翻译。

**表3-7　ICMPv6到ICMPv4消息类型的翻译**

| ICMPv6 消息类型 | | 翻译的ICMPv4消息类型 | |
| --- | --- | --- | --- |
| 回显请求（128）和回复（129） | | 回显（8）和回复（0） | |
| 多播监视器发现查询，报告，结束（130，131，132） | | 通常本地连接，抛弃 | |
| 邻居发现消息（133～137） | | 通常本地连接，抛弃 | |
| 目的地不可达（1） | Code 0（没有路径到达目的地） | 目的地不可达（3） | Code＝1（主机不可达） |
| | Code 1（与目的地的通信被禁止） | | Code＝10（与目的网络/主机的通信被禁止） |
| | Code 2（超过源地址范围） | | Code＝1（主机不可达） |
| | Code 3（地址不可达） | | Code＝1（主机不可达） |
| | Code 4（端口不可达） | | Code＝3（端口不可达） |
| 报文太大（2） | | 目的地不可达（3），Code＝4（需要分段数据报，设置DF位） | |
| 超时（3） | | 超时（11），有相同的代码 | |
| 参数问题（4） | Code 0（遇到错误头字段） | 参数问题（12） | Code＝0（遇到错误头字段）并且更新指针值或抛弃（按表3-8） |
| | Code 1（遇到无法识别下个报头类型） | | 目的地不可达（3），Code＝2（协议不可达） |
| | Code 2（遇到无法识别IPv6选项） | | 抛弃 |
| 其他代码或未知的ICMPv6类型 | | 抛弃 | |

对于参数问题的错误信息，报头指针值的翻译见表 3-8。

**表 3-8 ICMPv6 到 ICMPv4 报头指针值的翻译**

| ICMPv6 指针值 | 翻译的 ICMPv4 指针值 |
|---|---|
| 版本/流量类型（0） | 版本/IP 报文头长度（0） |
| 流量类型/流标签（1） | 服务类型（1） |
| 流标签（2，3） | 抛弃 |
| 负载长度（4，5） | 总长度（2） |
| 下一个报头（6） | 协议（9） |
| 跳数限制（7） | 生存时间（8） |
| 源地址（8 ~ 23） | 源地址（12） |
| 目的地址（24 ~ 39） | 目的地址（16） |

ICMPv4 校验和的计算也必须在进行完整的翻译过程之后。ICMP 错误的有效负载可能会导致翻译问题，特别是包含了另一种协议的 IP 地址。如果附加的 IP 报头也是返回的错误信息的一部分，翻译网关通常会企图翻译这个附加的 IP 报头，如果 IPv6 地址超出了分配的可转换的 IPv4 地址空间，就会无法实现。

### 3.3.1.4 IP 报头翻译

IP 报头的翻译过程应用于每个数据报以下字段映射。双向翻译的字段映射都总结如下：

| IPv4→IPv6 报头翻译 | IPv6→IPv4 报头翻译 |
|---|---|
| 版本 = 6 | 版本 = 4<br>报头长度 = 5（没有 IPv4 选项） |
| 流量类型 = IPv4 报头 TOS 位数或翻译器配置值 | 服务类型 = IPv6 报头流量类型字段或翻译器配置值 |
| 流标签 = 0 | |
| 负载长度 = IPv4 报头总长度 -（IPv4 报头长度 + IPv4 选项长度） | 总长度 = IPv6 报头负载长度 + IPv4 报头长度<br>id = 0 |
| 下个报头 = IPv4 报头协议字段值［将 ICMP（1）改成 ICMPv6（58）］ | 标识 = 不要分段 = 1，更多分段 = 0（除非 IPv6 数据报有一个分段的报文头表明允许分段）<br>分段位移 = 0 |
| 跳数限制 = IPv4 生存时间 - 1 | 生存时间 = IPv6 跳数限制 - 1<br>协议 = IPv6 下个报头字段；ICMPv6（58）改成 ICMP（1）并且 IPv6 报头逐跳（0），IPv6 路由（43），IPv6 标志，还有 IPv6 选项（60）由于不适合 IPv4 被舍弃 |

（续）

| IPv4→IPv6 报头翻译 | IPv6→IPv4 报文头翻译 |
| --- | --- |
| | 报头校验和 = 通过最新的 IPv4 报头计算 |
| 源 IP 地址 = 相关 IPv6 前缀和源 IPv4 地址基础上的 IPv4 可翻译的 IPv6 地址 | 源 IP 地址 = 从 IPv4 可翻译的 IPv6 地址推导的 IPv4 地址，这个 IPv6 地址落在 IPv6 可翻译的前缀；或映射的 IPv6 地址，这个 IPv6 地址基于翻译器的有状态地址映射（绑定信息库） |
| 目的 IP 地址 = 从目的 IPv4 地址（无状态）推导的 IPv4 可翻译的 IPv6 地址；或映射的 IPv6 地址，这个 IPv6 地址基于翻译器的有状态地址映射（绑定信息库） | 目的 IP 地址 = IPv4 转换的 IPv6 目的地址的 IPv4 部分<br>选项 = 无 |

现在来看看一些利用了 IP/ICMP 翻译算法的用于翻译 IPv4 和 IPv6 数据报的技术。

### 3.3.2 主机泵

主机泵（Bump In the Host，BIH）是一种基于主机的 IPv4/IPv6 翻译技术，让一台运行 IPv4 应用的主机可以与 IPv6-only 的主机通信。概念是在任何底层的 IPv6 通信的技术中屏蔽 IPv4 应用。BIH 应用的 IPv4 应用类别，包括那些使用 DNS 的地址解析和不在应用协议数据域中使用 IP 地址的应用。RFC 6535（即本书参考文献［49］）中指出，BIH 不推荐在与 DNS64 的连接中使用，它会导致双重协议转换，只有在原生双协议栈或隧道不能使用的时候推荐使用。

BIH 是堆栈泵（Bump In the Stack，BIS[50]）技术和 API 泵（Bump In the API，BIA[51]）技术组合的继任者。同样的，它既可以像 BIS 一样在网络层转换 IP 数据报，也可以像 BIA 一样在应用程序接口或套接字层中转换。

API（套接字）层策略，有两种架构，可二择其一，在 IPv4 和 IPv6 的 API 之间转换，或在应用和主机上的传输层栈中执行。BIH 基于套接字的体系结构如图 3-21 所示，由 API 转换器、地址映射器、扩展名解析器和函数映射器组成。

当 IPv4 应用发送 DNS 查询目的主机的 IP 地址时，扩展名解析器会拦截该查询并且生成一个额外的要求 AAAA 记录的查询。A 记录查询的 DNS 回复会提供给定的 IPv4 地址的 API 查询的答案。由于仅有 AAAA 记录的解析，这使扩展名解析器从地址映射器请求 IPv4 地址。域名解析器利用映射的 IPv4 地址产生 A 记录通过 API 回复应用。地址映射器维护 IPv6 地址到那些从私有地址空间组成的内部地址池的 IPv4 地址的映射。函数映射器拦截 API 函数的调用，并且匹配 IPv4 API 调用到 IPv6API 调用，然后返回 IPv4 API 调用回复作为结果。

网络层的方法侦查 TCP/IPv4 模块和连接层设备（如网卡）之间的数据流，并且将 IPv4 数据报转换成 IPv6 格式。BIH 网络层体系结构如图 3-22 所示。

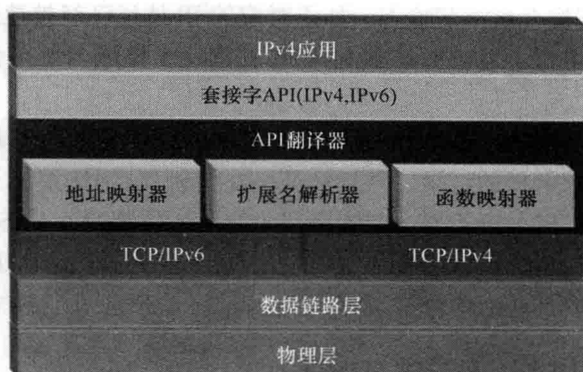



图 3-21　BIH 基于套接字的体系结构

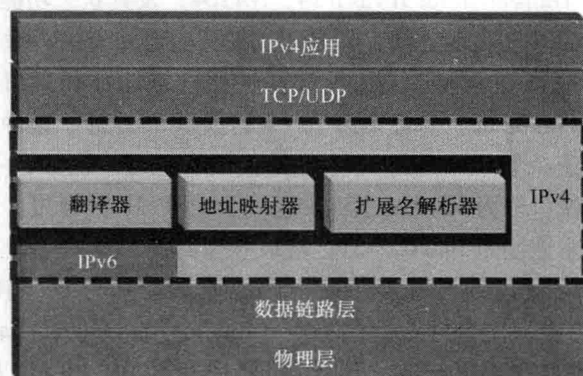



图 3-22　BIH 网络层体系结构

翻译器组件根据前面描述的 IP/ICMP 翻译算法，将 IPv4 报头转换成 IPv6 报头。扩展名解析器侦听 A 记录类型的 DNS 查询；一旦侦听到这样一个查询，扩展名解析器组件就产生一个 AAAA 记录类型的查询给该主机域名（Qname）和级别（Qclass）。如果没有从 AAAA 记录查询收到肯定的回复，接下来的通信就会使用 IPv4；如果 AAAA 记录查询被成功解析，扩展名解析器就会命令地址映射器组件将返回的 IPv4 地址（A 记录）和返回的 IPv6 地址（AAAA 记录）关联起来。如果仅收到 AAAA 记录的响应，地址映射器就会从私有 IPv4 地址的内部地址池中分配 IPv4 地址。

由于需要一个 IPv4 地址为应用请求 A 记录查询的结果提供响应。因此，地址映射器维护着真实或手动分配的 IPv4 地址与目的地的 IPv6 地址之间的对应关系。任何定位到该 IPv4 地址的数据报将会被翻译器转换成 IPv6 数据报通过 IPv6 网络传输。

映射一个给定的 IP 地址到主机域名的 PTR 记录请求可采用 BIH 的方式处

理。PTR 调用/查询会被拦截检查，如果对应的 IP 地址已经被地址映射器映射的话，就会产生一个对于对应 IPv6 地址 PTR 查询，主机域名结果会被关联到初始的请求。

BIH 的套接字版本，对 DNS 查询的 DNSSEC 验证是原生支持的，这是因为套接字调用仅请求解析的结果，并且验证是在解析器/网络层面上处理。网络层上的支持需要在扩展名解析器上配置密钥才能保证能够验证 DNSSEC 的响应。

如果从没有被地址映射的外部主机发出的 IPv6 数据报被 BIH 主机接收，地址映射器就会从内部地址池分配 IPv4 地址并且将 IPv6 报头转换成 IPv4。

### 3.3.3　IPv6/IPv4 的网络地址翻译

RFC 6146（即本书参考文献［52］）中定义了 IPv6/IPv4 的网络地址翻译（Network Address Translation for IPv6/IPv4，NAT64）及其状态功能操作。NAT64 能让 IPv6 主机向 IPv4 主机发起连接，但是不支持反向建立连接，除非 NAT64 网关中已经配置好 IPv4/IPv6 的地址映射。NAT64 使用网络地址和端口转换方法（Network Address and Port Translation，NAPT）来转换 IPv4 地址。这种方法能通过区分 TCP/UDP 端口号让一个 IPv4 地址映射到多个 IPv6 地址。例如，一个 IPv6 地址为 2001：db8：：1、端口号为 4040 的主机发送一个 UDP/IP 数据报，可能会被 NAT64 的网关映射到 IPv4 地址为 192.0.2.31、端口号为 1024，而另一个 IPv6 主机使用地址 2001：db8：：2、端口号 3701 则被映射到地址 192.0.2.31、端口号 1025。这种协议映射信息被存储在一个绑定信息基地（Binding Information Base，BIB），其中的三种协议信息都是动态维护的：TCP、UDP 还有 ICMP（ICMP 标识符是关联地址而不是端口号的）。同样的三张会话表，对每一个协议都进行维护，根据 IPv4 或 IPv6 的源或目的地址和端口号追踪每个对话。如图 3-23 所示（对图 3-20 稍微做了些修改），例中添加了端口号，用 p =〈port〉在每一个数据报中表示。如果是一个 TCP 会话，那么 TCP BIB 就会包含：

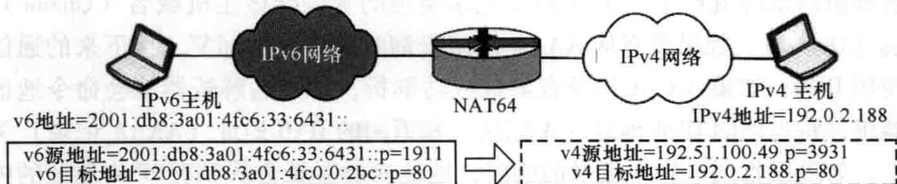



图 3-23　NAT64 协议和地址转换

(2001:db8:3a01:4fc6:33:6431 ::,1911)⇔(195:51:100:49,3931)

并且 TCP 会话也将会跟踪这个入口：

(2001:db8:3a01:4fc6:33:6431 ::,1911),(2001:db8:3a01:4fc0:0:2bc ::,

80)⇔(195:51:100:49,3931),(192:0:2:188;,80)

因此在这基础上，NAT64 执行两个功能：根据前面章节讨论的进行 IP/ICMP 协议转换和地址转换去映射入站和出站的地址。地址转换需要 NAT64 网关维护两个地址池：一个是在 IPv6 网络用 IPv6 地址池表示 IPv4 地址，另一个是在 IPv4 网络用 IPv4 地址池表示 IPv6 地址。IPv6 地址池由前面讨论的 IP/ICMP 转换分配的前缀组成，如对于前面讨论的例子为：2001：db8：3a01：4f00：：/56。IPv4 地址池是公有 IPv4 地址的分配，这些地址被用来为 IPv6 主机创建 IPv4 连接，前面的例子则是 195.51.100.0/24。接下来将讨论 NAT64 和 DNS64 之间的关系。

### 3.3.3.1　NAT64 和 DNS64

正如刚才讨论到的，NAT64 使用本章前面介绍的 IP/ICMP 翻译手段将 IPv6 数据报翻译为 IPv4 数据报，通过添加可选的有状态组件，可实现与被转换的 IPv4 地址无关的 IPv4/IPv6 映射处理。使用无状态或者有状态的翻译可以使 IPv6 主机与 IPv4 目的地通信。

这个策略的关键是 DNS64 组件，它是一个特殊的递归 DNS。通常，它执行 AAAA 记录的查询并且传递 IPv6 地址的有效响应；但是，当 AAAA 记录的响应失败的话，它会执行 A 记录的查询，以便在缺失有效 IPv6 地址的时候，能试图识别出 IPv4 目的地址。如果一个有效的 A 资源记录集被 DNS64 接收，它就会用与前面描述过的 IP/ICMP 翻译算法类似的方法表示出包含 IPv4 转换的 IPv6 地址的 AAAA 查询结果，并且发送到解析器。DNS64 解析过程如图 3-24 所示。

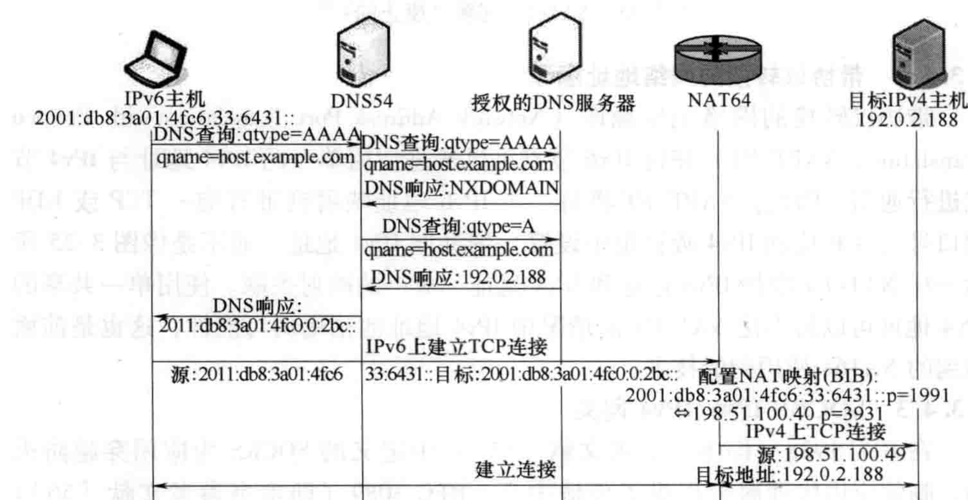



图 3-24　DNS64 解析过程

## 3.3.4 其他翻译技术

为了完整起见和某些历史问题，在这里总结了其他利用网关的翻译技术。

### 3.3.4.1 带端口转换的网络地址翻译

顾名思义，带端口转换的网络地址翻译（Network Address Translation with Port Translation，NAT-PT）[53]会将 IPv4 地址翻译成 IPv6 地址（类似 IPv4 NAT），而且还有前面所说的协议报头翻译的功能㊀。NAT-PT 设备在 IPv6 网络和 IPv4 网络之间充当网关的角色。例如，可以能够使原生的 IPv6 设备与 IPv4 网络上的主机通信。NAT-PT 设备维护着 IPv4 地址池，当通信发生的时候，它会分配一个已有的 IPv4 地址到一个 IPv6 地址上。图 3-25 给出了 NAT-PT 部署（废止的）。由于 RFC 4966（即本书参考文献［54］）里提及的众多理由，NAT-PT 已经被废止并且不再部署使用了。

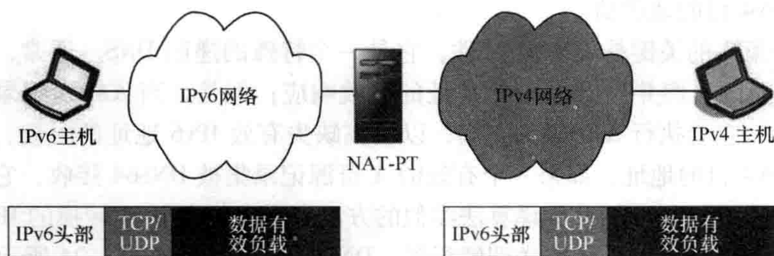



图 3-25 NAT-PT 部署（废止的）[53]

### 3.3.4.2 带协议转换的网络地址翻译

带协议转换的网络地址翻译（Network Address Port Translation with Protoco Translation，NAPT-PT）使得 IPv6 节点可以通过使用单一的 IPv4 地址与 IPv4 节点进行通信。因此，NAPT-PT 将每一个 IPv6 地址映射到带有唯一 TCP 或 UDP 端口号（在相应的 IPv4 数据报中设置）的共同 IPv4 地址，而不是像图 3-25 所示一样 NAT-PT 维护 IPv6 地址和 IPv4 地址一对一的映射关联。使用单一共享的 IPv4 地址可以最小化 NAT-PT 的情况中 IPv4 地址池枯竭的可能性，这也是前面提到的 NAT64 使用到的技术。

### 3.3.4.3 SOCKS IPv6/IPv4 网关

在 RFC 1928（即本书参考文献［55］）中定义的 SOCKS 为应用穿越防火墙、高效应用代理服务提供了传输中介。RFC 3089（即本书参考文献［56］）应用了 SOCKS 协议来进行 IPv4 和 IPv6 之间的翻译。类似前面讨论过的几种翻译

㊀ NAT-PT 网关之间关联所管理的源和目标 IP 地址字段除外。

技术，这个方法包括特殊的 DNS 处理，被称为“DNS 解析委托”，其将域名解析从解析客户端委托到 SOCKS IPv6/IPv4 网关。IPv4 或 IPv6 的应用可以在“SOCKS”化后与 SOCKS 网关代理通信，与使用不同协议的主机建立唯一连接。图 3-26 所示的基本的 SOCKS 网关配置，阐明了配置了 SOCKS 客户端的 IPv6 主机连接 IPv4 主机的情况。经过“SOCKS 化”的 IPv4 主机也可以通过 SOCKS 网关连接到 IPv6 主机，连接方式如图从右到左所示。

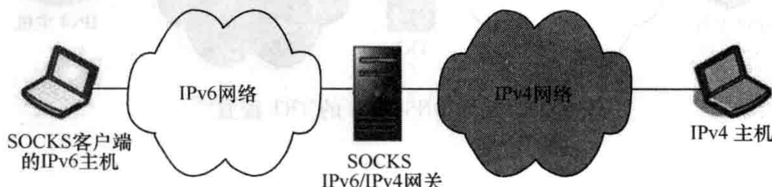



图 3-26 基本的 SOCKS 网关配置[56]

#### 3.3.4.4 传输中继翻译器

传输中继翻译器（Transport Relay Translator，TRT）[57] 是一个十分类似 SOCKS 的配置，TRT 的特点是使用一个有状态网关将两个独立连接通过不同网络连接起来。TCP 和 UDP 的连接从主机开始到 TRT 终止，然后 TRT 新建一个单独的连接到目的主机，并且这个连接成为两个连接之间的中继。TRT 需要一个 DNS 应用层网关（DNS-ALG）㊀，DNS-ALG 起到 DNS 代理的作用。TRT 可以让 IPv6 主机与 IPv4 目的地址进行通信。因此，DNS-ALG 的主要功能是根据 IPv6 解析器的要求查询一条 AAAA 记录；如果记录被返回，回复将通过解析器、数据连接和 IPv6 连接。如果没有 AAAA 记录返回，DNS-ALG 将会执行一条 A 记录的查询；如果收到回复，DNS-ALG 将会用包含在 A 记录里面的 IPv4 地址生成一个 IPv6 地址。RFC 3142 不仅定义了 TRT，而且还规定了使用前缀必须要在 C6::/64 后跟着 32 位 0 再加上 32 位的 IPv4 地址。然而，INNA 并没有分配 C6::/64 前缀。因此，需要配置一个本地前缀（见图 3-27）。

#### 3.3.4.5 应用层网关

应用层网关（Application Layer Gateway，ALG）在应用层中执行协议转换和应用代理功能，类似 HTTP 代理。客户端应用如果配置好服务器代理的 IP 地址，就可以通过打开该应用建立连接，如 web 浏览器中的 HTTP 代理。如果是通过 IPv6 网络上的主机连接到 IPv4 网络的话，ALG 无论对于 web 还是特定应用程序的访问都是相当有用的。

---

㊀ 有时被称为“不给糖就捣蛋的 DNS-ALG”或 totd。

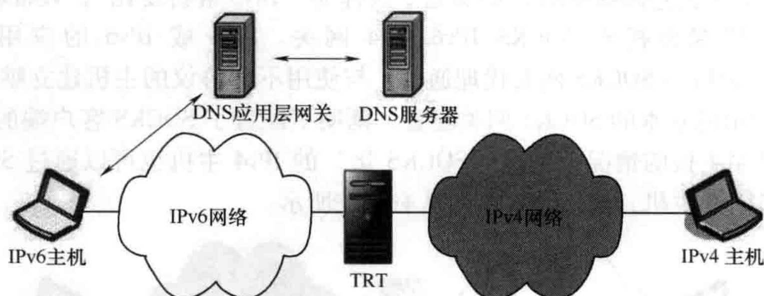



图 3-27 使用 DNS-ALG 的 TRT 配置[57]

## 3.4 IPv6 的应用支持

实际上，TCP/IP 应用的应用程序接口（Application Programming Interface，API）是套接字接口，它最早是在 BSD UNIX 系统上实现的。套接字定义了程序调用让应用通过 TCP/IP 层的接口在 IP 网络上交换信息。微软的 Winsock API 也是以套接字为基础的。为了支持 IPv6 更大的地址空间和附加特性，sockets 和 Winsock 接口都已经过修改了。事实上，大部分主流的操作系统都已经实现了对 sockets 或者 Winsock 的支持，包括微软 Windows（XP SP1，Vista，7，Server 2003 & 2008）、Solaris（8 +）、Linux（kernel 2.4 +）、Mac OS（X.10.2）、AIX（4.3 +）和 HP-UX（11i with upgrade）操作系统。

更新后的套接字接口不仅支持 IPv4 和 IPv6，而且通过使用 IPv4 映射的 IPv6 地址实现了 IPv6 应用与 IPv4 应用的交互操作。检查应用厂商对 IPv6 的兼容性和要求，如果你的程序使用的是没经过更新的套接字的程序调用，如 gethostbyname ()，那就说明你必须要对这部分的应用进行更新了。更多的细节可以参考本书参考文献［58］给出的应用程序开发者经验。

## 3.5 服务提供商的 IPv4/IPv6 共存⊖

服务提供商，特别是家庭宽带服务提供商，可以在它们已有的网络上实现 IPv6 并最终给客户分配 IPv6 地址。然而，它们的网络一般来说比企业网络更为复杂，企业网络可以大致分为企业内部网络域和外部面向因特网的公共域。站在同一个高度，也许从简化的网络视图可以发现，服务提供商给这个两域模型

⊖ 该部分内容是基于本书参考文献［59］。

增加了第三个域，即客户接入网络。增加这第三个域需要利用一个包含一个或多个到本章为止所讨论过的技术的 IPv6 部署方法。

用户驻地设备（Customer Premises Equipment，CPE），通常是一个路由器、电缆调制解调器、光纤终端单元或无线路由设备，这些使服务提供商的接入链路终止的这一类设备，被称为用户边缘（Customer Edge，CE）路由器。CE 路由器通过提供商边缘（Provider Edge，PE）路由器将所有客户发送的 IP 数据报转发到服务商网络。PE 路由器是服务提供商网络连接终止的一端。PE 路由器会将数据报路由到其他面向用户的 PE 路由器，或直接路由到因特网，或通过服务提供商的核心或骨干路由器转发（也被称为提供商或 P 路由器）。

相反，PE 路由器，将因特网或其他商业网络应用产生的服务流量，从提供商网络，以输入流量路由到用户站点。服务提供商的“核心”网络由在其自身及 PE 路由器之间路由，数据报的骨干路由器组成。服务提供商一般为 CE 设备面向服务提供商的网络接口提供 IP 地址。对于商业应用，CE 路由器是在相应的客户网站中到其他一些网络的接口；而对于家庭型应用，CE 设备有时是一个 DHCP 服务器，为相对较少数量的主机提供 IP 地址。

### 3.5.1　参考架构

在服务提供商的网络中，IP 数据报可以通过各种基本服务，如 MPLS，最终被路由到目的地。对于需要通过公共网络来进行访问的目的地址，目的主机应该能够通过 IPv4 地址（通过“IPv4 因特网”）或 IPv6 地址（通过“IPv6 因特网”）或以上两者来进行访问。尽管这里将其说明为两个逻辑上独立的网络，根据所使用 IP 路由的版本来区分，但在物理上这是同一个网络。连接到一个或两个这样的“因特网”取决于服务提供商的能力及路由通告在 IPv4、IPv6 或两者上同时被支持情况（见图 3-28）。

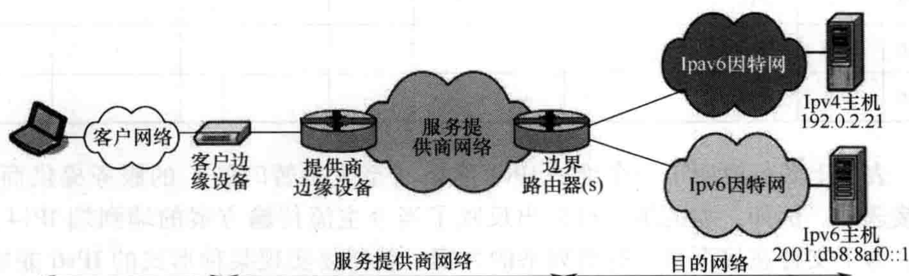



图 3-28　基本三层架构：客户网络/服务提供商网络/目的网络

下面将对基本三层架构中的每一层来考虑 IPv6 的部署：

- 用户网络。取决于服务提供商对用户是否支持 IPv4、IPv6 或两者都支持。

• 服务提供商网络。服务提供商网络的边缘设备及“核心”设备，将用户与因特网相互连接。

• 目的网络。取决于目的网站、电子邮件服务器等的能力，可达性可能需要支持一个或两个 IP 版本

### 3.5.2　部署方法概述

服务提供商通常控制它们的网络，在大多数情况下是控制如何将 IP 地址分配给 CE 设备中面向服务提供商的接口。客户可以自行实现任一 IP 版本，尽管它们通过特定 IP 版本的连接能力，在一定程度上取决于预期目的地对协议版本的支持和服务提供商对传输的支持。表 3-9 列出了基于所支持的协议版本的基本部署选项，其中总结了在这个简单的三层架构中每层使用不同 IP 版本时的可用连接选项。

**表 3-9　基于所支持的协议版本的基本部署选项**

| 客户网路 | 服务提供商网络 | 目的网络 | IPv6 部署方法 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NAT44 | 双栈 | 6PE/6VPE | 配置隧道 | 6rd | 4over6 | 轻型双栈 | 具有 DNS64 的 NAT64 | 全 IPv6 |
| IPv4 | IPv4 | IPv4 | · | · | | | | | | | |
| IPv4 | IPv4 | IPv6 | | · | | | | | | | |
| IPv6 | IPv4 | IPv4 | | · | | | | | | | |
| IPv6 | IPv4 | IPv6 | | · | · | · | · | | | | |
| IPv4 | IPv6 | IPv4 | | · | | · | | · | · | | |
| IPv4 | IPv6 | IPv6 | | · | | | | | | | |
| IPv6 | IPv6 | IPv4 | | · | | | | | | · | |
| IPv6 | IPv6 | IPv6 | | · | | · | | | · | | · |

表中上四行说明了一个维护 IPv4 网络（至少是暂时的）的服务提供商的连接选项。例如，表的第一行突出反映了当今主流传输方案的端到端 IPv4 连接。为了支持在接下来三行所列举的方案，就需要实现某种形式的 IPv6 兼容。在第二行及第三行所列举的情况下，为支持 IPv4 客户与 IPv6 目的地的通信或相反方向的通信，需要部署双协议栈。第四行表明通过服务提供商的 IPv4 网络连接两个 IPv6 端点。很多种部署方法可以解决这种情形，包括双协议栈、端到端配置隧道（如各种 VPN）、6rd；或如果使用 MPLS 的话，6PE 及 6VPE

也能胜任。

表中下四行说明一种 IPv6 的服务提供商的实现及支持多种连接类型的能力或要求。当然，服务提供商可以在一次阶段性的 IPv6 部署期间实现多种技术，如根据市场或地理位置进行分阶段实施。

从表中可以清晰地看到，双协议栈支持所有的组合方案，尽管其需要完备的 IPv4/IPv6 地址分配方案，以及要求 PE 设备或所有的服务提供商路由器部署双协议栈。应当注意的是，当面向互联网配置双协议栈或其他部署方法同时支持两个版本的协议时，边界路由器必须配置为对 IPv4 及 IPv6 路由路径进行通告。

## 3.5.3　路由基础设施的部署方法

本节对每种实施策略进行概述。一般不会通过配置隧道选项来解决问题，因为这意味着要一个事先配置好 VPN，或一个穿越服务提供商网络而不由服务提供商直接参与的连接客户和目的网络主机的隧道。自动隧道或“软线（softwires）”是快速创建的隧道，下面将看到它是几种部署方法的核心部件。

### 3.5.3.1　NAT444

NAT444[60]在技术上来说并不是一种 IPv6 实现策略。但是 NAT444 作为一种为延长 IPv4 生命周期“争取时间”的方法，其目的是为了部署 IPv6 形式的补充地址空间。NAT444 拥有可以作为大型（“运营商级别”）IPv4- IPv4 网络地址转换网关（LSN NAT44）的特点，这个特点使多个用户共享一个公共 IPv4 地址。NAT444 有不需要更换现存的 IPv4 网络用户端设备的好处，尽管需要付出限制用户会话数量（可能会被如 AJAX、RSS 订阅等应用所需要），用户端设备位置信息的丢失（E911），及其他更多[61]的代价（见图 3-29）。

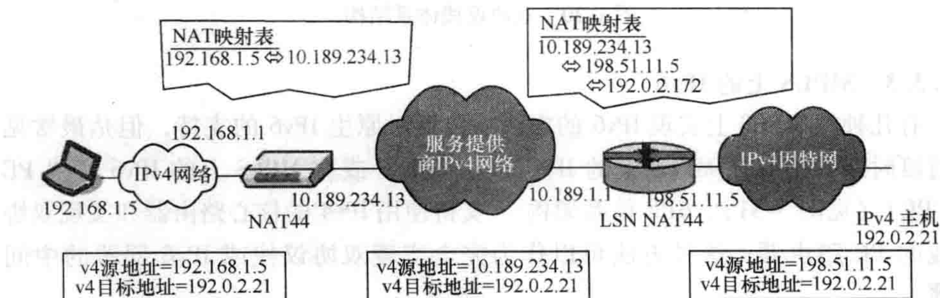



图 3-29　NAT444 体系结构

图 3-29 所示的 NAT444 体系结构，显示了在数据通路中使用了两个 NAT 的

情况：第一个在用户端设备中，将家庭网络中的 IPv4 地址空间转换为服务提供商提供的私人地址空间；第二个为 NAT，即 LSN NAT44，将用户的客户端地址转换为一个公共 IPv4 地址。不同的端口号用来在同一个或多个终端用户中区分不同的会话。术语“NAT444”是对使用两个 IPv4-IPv4 NAT 的描述，NAT444 实际是转换三个 IPv4 地址空间（用户私人地址、服务提供商接入地址和公用网络）的 IP 地址。

### 3.5.3.2 双协议栈

如前所述，双协议栈的实现需要对每个基础设施设备（或至少是路由设备，也可能是设备接口）进行 IPv4 地址及 IPv6 地址的配置。不管相应的服务提供商所支持的协议是 IPv4 还是 IPv6，支持双协议栈用户端设备都可以高效运行；然而端到端的连接则要求在服务提供商网络与相应的目的网络之间的协议连续性。

服务提供商网络中对双协议栈的支持可能是完全部署整个网络或只在网络“边缘”部署。例如，PE 路由器可以实现双协议栈来启用对 IPv4 和 IPv6 用户的支持，而面向因特网的端口则使用目的网络的 IP 版本建立连接（见图 3-30）。然而，如果没有在与 PE 路由器互联的骨干路由器上实现完全的双协议栈部署的话，则需要 PE 之间的隧道来传输未获实现的 IP 版本的流量。这个实现方法将在下一小节“MPLS 上的 IPv6 部署”中说明。

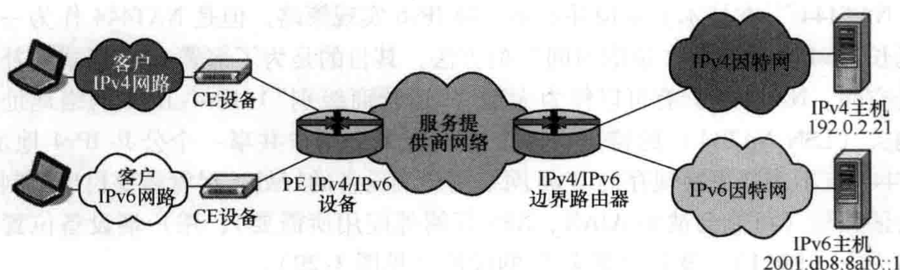



图 3-30 双协议栈体系结构

### 3.5.3.3 MPLS 上的 IPv6

有几种在 MPLS 上实现 IPv6 的方法，包括对原生 IPv6 的支持，但是最常见的过渡时期方法有在 MPLS 上的 IPv6 PE（6PE）或在 MPLS 上的 IPv6 VPN PE（6VPE）（见图 3-31）。6PE 体系架构[62]支持使用 IPv4 的核心路由器和实现双协议栈的 PE 路由器；这种方法可以作为完全实现双协议栈或 IPv6 部署的中间步骤。

双协议栈 PE 路由器通过在 IPv4 上的多协议边界网关协议（Multi-Protocol Border Gateway Protocol，MP-BGP）来传达其面向核心 IPv4 网络的 IPv6 地址可达性。这将使得入口 6PE 路由器能够识别出口 6PE 路由的 IPv4 地址、识别标签交

换路径（Label Switch Path，LSP）和相关联的 IPv4 标签，从而使得标签交换能够通过骨干 IPv4 路由器达到出口 6PE 路由器。这种技术需要使用一个 IPv6 标签和一个外部 IPv4 标签，但必须预先定义 IPv6 over IPv4 隧道。

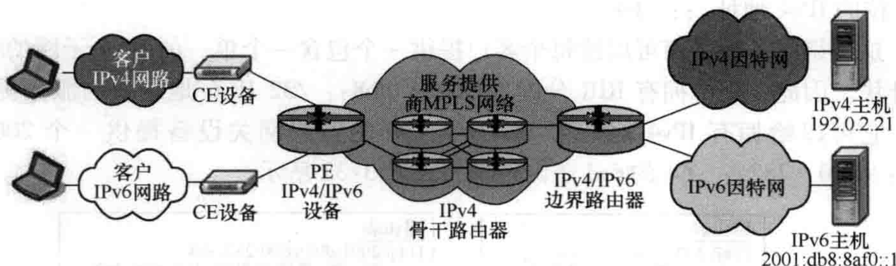



图 3-31　6PE 体系结构

6VPE[63] 架构与 6PE 是高度相似的，不同在于 6VPE 使用 IPv6 over IPv4 的"VPN 隧道"来穿透核心网络。相对于 6PE，6VPE 可以提供更好的隐私保护和地址空间复用。6VPE 将给定的客户 VPN 与 PE 设备中的一个或多个 VPN 路由和转发（VPN Routing and Forwarding，VRF）表项相关联。PE 路由器会将用户网络的物理链路信息（可能是第 2 层的报头信息）与 VRF 表中对应的 VPN 关联在一起。

每个 CE 设备向其相连的 PE 路由器发送路由通告（相应网站的可达性）。MPLS 标签会被分配给 VPN（通过 VPN、CE 设备或者路由来分配），当 MP-BGP 路由信息在服务提供商网络中面向用户网络（VPN）的 PE 路由器之间分发的时候，这个 MPLS 标签也会在路由路径上被传递。IPv6 路由路径在 VPNv6 地址族中的 MP-BGP 信息中承载，其中的下一跳可达性信息被配置为 IPv4 映射的 IPv6 地址［::ffff: <ipv4 地址>］。由于骨干网只使用 IPv4，按照 BGP 要求下一跳地址应当是在同一地址族内，因此下一跳地址也是一个 IPv4 地址。当 IPv4 分组从一个 CE 路由器到达一个 PE 接口时，PE 路由器从 VRF 表中确定 VPN，然后使用相应的标签将分组转换到适当的 PE 设备并最终到底目的地。

### 3.5.3.4　6rd（IPv6 快速部署）

RFC 5569（即本书参考文献［64］）将"在 IPv4 基础设施上快速部署 IPv6（6rd）"定义为一种在维持 IPv4 基础设施的情况下使服务提供商能够给终端客户实现 IPv6 连接和提供 IPv6 地址的技术。RFC 5969（即本书参考文献［65］）为 6rd 协议作出了规范。这种方法要求通过一个改进过的 6to4 技术将客户的 IPv6 数据流量从客户端通过"软线隧道"（softwire tunneling）送达至一个 IPv6 目的地。这个改进要求用服务提供商的 IPv6 前缀（/32）代替 6to4 的前缀，相对于依靠松散维护的 6to4 任播中继，使用 2002::/16 提供了一个更好的双向控制。

就像6to4那样，6rd的IPv6前缀后的32位由6to4网关的IPv4地址组成，在前文这种情况下是客户端宽带路由的地址。因此，6to4的前缀定义为2002：{32位的IPv4地址}：：/48，而6rd的前缀定义为{32位的服务提供商IPv6前缀}：{32位的IPv4地址}：：/64。

这使得服务提供商可以给每个客户提供一个包含一个单一的IPv6子网的/64地址块。因此，一个拥有RIR分配的2001：db8：：/32 IPv6地址块的服务提供商，它可以给拥有IPv4地址为192.0.2.130的客户网关设备提供一个2001：db8：c000：282：：/64的6rd子网地址，如图3-32所示。

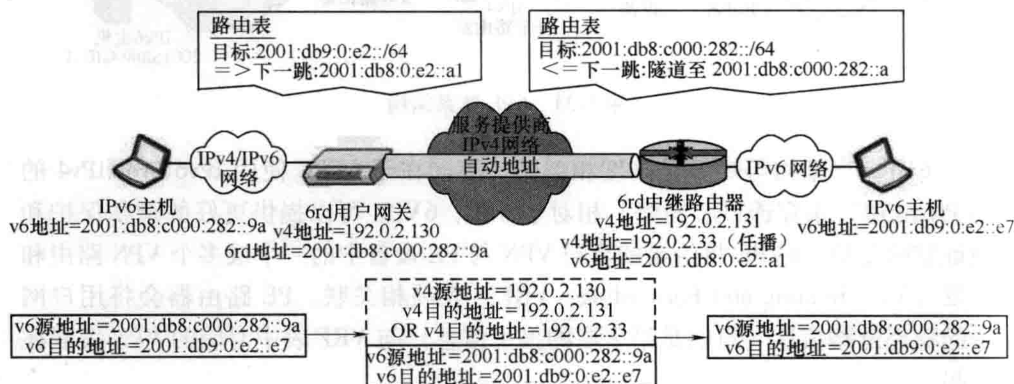



图3-32 6rd部署实例[66]

家庭设备所需要的IPv6地址将从这个子网中分配得到。例如，如图3-32所示，一台计算机被分配了IPv6地址2001：db8：c000：282：：9a。6rd客户网关通过隧道将从IPv4产生的IPv6分组传输到6rd网关（中继路由器）。6rd与6to4的另一个与地址有关的差异是，6to4的任播地址是固定的（192.88.99.1），而6rd的任播地址是由服务提供商根据自己的地址空间所定义的。必须给每个客户路由器提供6rd中继代理或任播地址。

6rd的中继路由器终止IPv4隧道的传输，然后将IPv6分组路由到其目的地。使用服务提供商的前缀，可使得6rd可达目的地址能够随同服务提供商的本地IPv6流量一同被通告。

### 3.5.3.5 4over6

4over6方法，在RFC 5747（即本书参考文献［67］）中被规范，是一种为了使IPv4用户通过一个IPv6网络到达IPv4目的地的自动隧道方法（如softwire）。作为一种与6PE相反的方法，4over6面对的是IPv6核心的骨干路由器，它对原生IPv6流量进行路由，对IPv4数据报则通过隧道传输。一个拥有IPv4地址空间的用户可以使用这种技术与IPv4目的地进行通信。

每个 CE 路由器对其相连的 PE 路由器提供路由更新。PE 路由器使用 MP-BGP 来传递 4over6 路由信息，并对网络流量进行相应的路由选择。就图 3-33 所示而言，当目的地 CE 设备（没有在图中标出）向 192.0.2.0/24 通告其可达性时，图中左侧的 CE 设备也向 198.51.100.0/24 通告可达性。通信双方的 PE 路由器都使用 MP-BGP 来传达可达性。当一个分组到达在图中左边的 PE 路由器时，其到达目的地的路由路径被标识。而且，为了经由 IPv6 核心（P）路由器路由，该分组需要使用 IPv6 报头封装起来。当出口的 PE 路由器收到 IPv6 分组时，这个 PE 路由器对分组进行解封装（移除 IPv6 报头），然后通过 CE 路由器将原来的 IPv4 分组路由到其目的地。

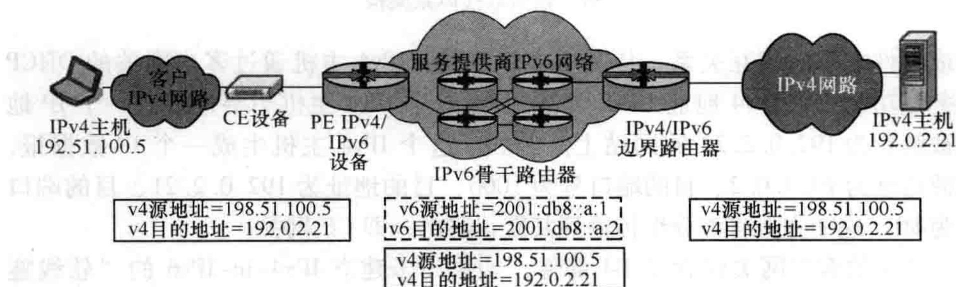



图 3-33　4over6 实例

当前的 4over6 架构只支持单个自治系统（Autonomous System，AS），因此，对多客户私人网络的支持是有限的，但是对多 AS 的支持是未来研究的方向。

### 3.5.3.6　轻型双栈

轻型双栈[68]是一种使得服务提供商能够在其网络中部署 IPv6 的技术，同时它也能使得对分配给用户网络设备的 IPv4 地址的长期支持与有效利用变得容易。服务提供商通常为客户路由器或网关这类接口直接指向宽带接入网络的设备分配 IP 地址。客户网关在给家庭网络中的 IP 设备分配 IP 地址时起到 DHCP 服务器的功能。这里假设这样的家庭网络设备将在相当长的一段时间内只支持 IPv4。

实现轻型双栈需要包含以下组件：

- 基本桥接宽带（Basic Bridging Broad Band，B4）部件。将 IPv4 家庭网络与 IPv6 网络桥接起来；B4 功能可能局限于客户网关设备内或服务提供商网络内。
- 软线 IPv4-in-IPv6 隧道。在 IPv6 上将 IPv4 流量在 B4 与地址族转换路由（Address Family Translation Router，AFTR）之间用隧道传输。
- AFTR。和 B4 部件一样作为 IPv4-in-IPv6 软线隧道的终止点，也执行 IPv4-IPv4 NAT 功能。

图 3-34 所示的轻型双栈体系架构，说明了在一个端到端 IP 连接中，三个

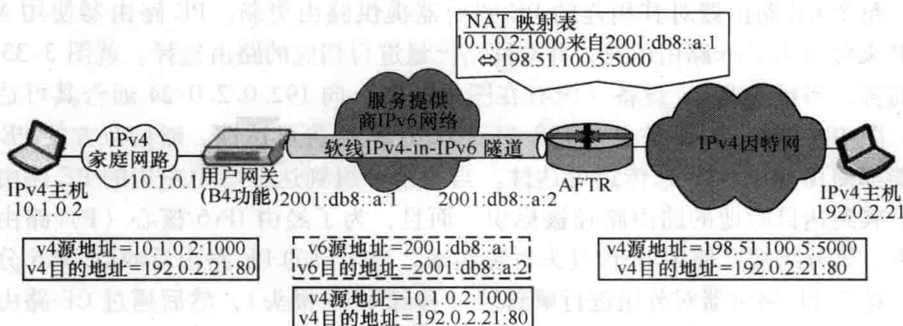



图 3-34 轻型双栈体系架构[68]

组成部件之间的相互关系。从图中左侧开始，IPv4 主机通过客户网关的 DHCP 服务器功能获取 IPv4 地址 10.1.0.2。假设这个 IPv4 主机想要连接到一个 IP 地址被解析为 192.0.2.21 的网站上。例如，这个 IPv4 主机生成一个 IP 数据报，其源地址为 10.1.0.2、目的端口号为 1000，目的地址为 192.0.2.21、目的端口号为 80。这个主机将该分组传输到其默认路由，即 CE 网关。

图中的客户网关包含了 B4 部件，如果尚未建立 IPv4-in-IPv6 的“软线隧道”，B4 会建立隧道。客户网关的 WAN 端口（面向服务提供商）已经被分配了一个 IPv6 地址，且在这个连接上隧道已经被建立起来。客户网关已经手动或者通过 DHCPv6 配置了 AFTR IPv6 地址。如图 3-34 所示，B4 部件将原本的 IPv4 分组用一个 IPv6 报头封装起来，然后将其传送到 AFTR。

AFTR 终止隧道并删掉 IPv6 头部。AFTR 随后行使 IPv4-IPv4 NAT 的功能。这需要将原本分组的私人（RFC 1918）IPv4 源地址转化为公共的 IPv4 地址。因此，在这种情况下，服务提供商必须为发往 IPv4 目的地分组的源 IP 地址，提供一个公共 IPv4 地址池。这公共地址池使得服务供应商能够更有效地利用日益稀缺的公共 IPv4 地址空间。AFTR 一般也进行端口转换，且为了双向正确映射 IPv4 地址和端口号，必须为每个 NAT 操作追踪这些映射。

图 3-34 中，AFTR 将客户的源 IPv4 地址及端口从 10.1.0.2：1000 映射为 198.51.100.5：5000。客户一般使用私人地址空间，这可能会发生地址重叠的现象，因此 NAT 映射表也追踪分组源自的隧道。包含 IPv4 地址及端口号为 198.51.100.5：5000 的数据报最终被传送到目的主机。去往这个地址/端口的返回分组被映射为目的地址为 10.1.0.2：1000，且通过隧道传输到 2001：db8::a：1。

部署 IPv6 或双协议栈主机的客户可以通过在客户网关中实现的 DHCPv6 功能或自动配置获得 IPv6 地址。从家庭网络传输到客户网关的 IPv6 分组不会使用“软线隧道”，而会被本地服务提供商的 IPv6 接入网络进行路由和传输。

#### 3.5.3.7 NAT64/DNS64

正如本章前面详细讨论过的，NAT64/DNS64 解决方案使得 IPv6 主机与 IPv4 目的地之间的通信变得简单。对服务提供商而言，这种方法可用在当客户端设备只支持 IPv6 的情况，因为这允许客户能与仍将存在相当长的一段时间的 IPv4 因特网进行通信。从服务提供商的角度来看，其基本架构如图 3-35 所示。一个 IPv6 主机需要发起一个 Web 会话时，可能为相应的 IP 地址发起 DNS 解析。在它们的递归域名服务器上部署了 DNS64 功能后，服务提供商能够自动将请求的主机地址（在这种情况下是 IPv4）转换为 IPv6 地址，来使得客户能够通过 NAT64 网关与目的主机建立 http 会话。

无论服务提供商何时为用户实现 IPv6，DNS64/NAT64 的部署能够使得这些用户访问终将会成为历史遗产的 IPv4 网站。

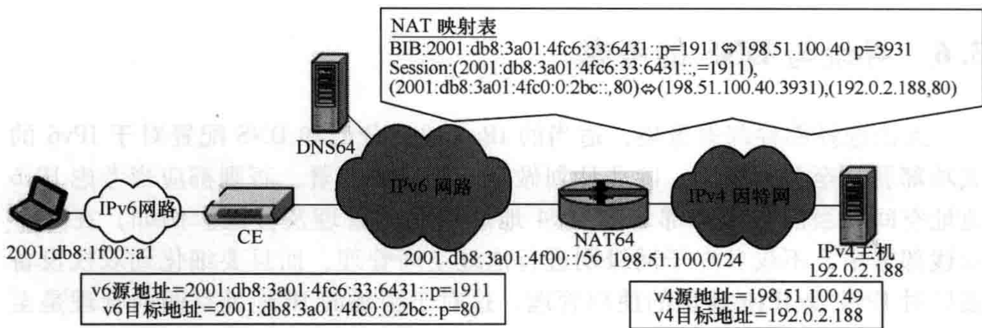



图 3-35 NAT64/DNS64 服务提供者架构

### 3.5.4 部署方法的比较

表 3-10 给出的部署方法的高层次比较，归纳了各部署方法的相关特性。

表 3-10 部署方法的高层次比较

| 基本标准 | IPv6 部署方法 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | NAT444 | 双栈 | 6PE/6VPE | 配置隧道 | 6rd | 4over6 | 轻型双栈 | 具有 DNS64 的 NAT64 | 全 IPv6 |
| 商业或住宅 | 均有 | 均有 | 商业 | 商业 | 住宅 | 商业 | 住宅 | 均有 | 均有 |
| 提供 IPv6 支持 | 不支持 | 支持 | 支持 | 支持 | 支持 | 支持 | 支持 | 支持 | 支持 |
| IPv4/IPv6 共存 | 不共存 | 共存 | 共存 | 共存 | 共存 | 共存 | 共存 | 共存 | 不共存 |
| 实现复杂度 | 高 | 高 | 中 | 中 | 低 | 低 | 中 | 高 | 中 |
| 需要客户端设备改变 | 不需要 | 不需要 | 不需要 | 不需要 | 需要 | 不需要 | 需要 | 不需要 | 需要 |

（续）

| 基本标准 | IPv6 部署方法 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | NAT444 | 双栈 | 6PE/6VPE | 配置隧道 | 6rd | 4over6 | 轻型双栈 | 具有 DNS64 的 NAT64 | 全 IPv6 |
| 需要新的服务提供商设备 | 需要 | 不需要 | 需要 | 不需要 | 需要 | 不需要 | 需要 | 需要 | 不需要 |
| 新进性操作和发现修理故障的复杂性 | 中 | 中 | 中 | 低 | 中 | 中 | 中 | 中 | 中 |
| 支持重量的 IPv4 地址空间 | 支持 | 不支持 | 6VPE | 不支持 | N/A | 不支持 | 支持 | 不支持 | N/A |

## 3.6　寻址与 DNS 的考虑

无论选择哪种部署策略，适当的 IPv6 地址分配和 DNS 配置对于 IPv6 的成功部署是至关重要的。除非计划做全新网络的部署，否则都应当考虑 IPv6 地址空间和当前与未来部署的 IPv4 地址空间的管理及分配。例如，在一个双栈部署中，不仅要在子网级别进行地址空间管理，而且要细化到双栈设备接口对 IPv4 及 IPv6 地址的使用管理，这对于精确的 IP 地址空间的管理是至关重要的。

要谨慎细致地分配从 RIR 获得的 IPv6 地址。作为一个一般原则，从 RIR 获得的地址块的前缀长度和准备分配给用户的地址空间大小之间的差异将决定需要如何处理地址的层级结构。例如，如果从 RIR 处获得一个/32 的地址块，并打算分配/64 的地址块给用户，那将有 32 位的地址空间可以用来进行层级分配，如通过区域（如/36s）、城市（如/44s）、服务节点（如/52s）、PE 设备（如/56s）这样的层次结构。例子中的层级分配结构将允许分配 16 个地区，每个地区又能分别分配 256 个城市，每个城市又能分配 256 个服务节点，每个服务节点又能分配 16 个 PE 设备，每个设备又能支持 256 /64s 个用户。当然，这是一个例子，可能分配更多或更少不同大小的层。

如果要分配更大的地址块给用户，那么将剩下更少的可使用位。如果是非均匀地将地址块分配给用户（如，/48 给企业用户，/64 给 SOHO 用户），那么在用稀疏、最适合或随机的方式分配地址空间时，就需要一个更加复杂的地址分配及追踪机制，这将会在本书第 5 章进行描述。

DNS 配置将使终端用户通达网络（命名空间）中 IPv4 或 IPv6 可寻址的目的地址。目的地址的解析属于各自域名管理员的管理范围。对于 IPv6 地址

（AAAA 记录查询类型）而不是 IPv4（A 记录查询类型）所返回的响应，将对查询主机表明这个地址只能通过 IPv6 进行访问。在这种情况下，IPv6 服务会被用来建立连接。

一些 IPv4/IPv6 转换技术对于 DNS 有直接的要求，如 NAT64/DNS64。如果给用户指定了反向区和地址空间，那么 DNS 必须要有正确（精确！）的 ip6. arpa 区及对应的 NS/glue 记录。这与分配 IPv4 地址的过程类似，但在十六进制与十进制的域名标签表示上有明显的语法差异。

# 第4章 IPv6准备情况评估

IPv4是如今使用最普遍的协议，它不仅是因特网的通信工具，同时作为广泛的应用工具服务于商业和家庭访问。迄今为止，IPv6还只是被使用在一个比较有限的范围。然而，这种情况正在改变，将来很可能会促进IPv6的使用，从而使IPv6的地位赶超其老大哥IPv4。为什么采用IPv6需要这么长的时间？这个问题也是在过去15年里人们一直在讨论的主题。答案就是并没有像电影《世纪骇客》那样的强迫性事件来逼迫人们需要在某个截止时间之前完成这技术上的升级。虽然IPv4地址空间在某个时刻已经被耗尽，但像NAT等技术的产生延长了IPv4的寿命，同时也拖延了IPv6的采用。

尽管包括美国政府在内的一些组织已经确定了它们IPv6部署的目标日期，IPv4/IPv6过渡技术仍需要支持很长一段时间的共存期和过渡期。但什么将会成为IPv6因特网的触发器，并推动IPv6占据主导地位呢？这又将在何时发生？移动设备的使用是否会促使IPv6增长？又或者使可用IPv4地址空间的消耗、云计算，还是数据中心虚拟化的激增呢？到底最终使IPv6具有更加突出的地位的引爆点是什么呢？答案是“以上都是”。

IPv6的广泛采用将会使这样一个逻辑过程得以实现，它以网络设备提供商支持IPv6开始，然后到服务提供商，再到应用/内容提供商，最后将其逐渐到企业和家庭使用。目前这个过程正在顺利进行，大部分设备提供商和服务提供商已经实现了IPv6。

2012年6月，美国Google（www.google.com）、Facebook（www.facebook.com）、Bing（www.bing.com）和Yahoo（www.yahoo.com）等服务提供商和内容提供商都永久性地在其主页上启用了IPv6。这次“世界IPv6启动”是一项重要的里程碑，标志着IPv6在因特网上被实际采用的开始。而这产生的结果就是，因特网协会报告有成千上万的公司和网站现在都支持IPv6了[69]。横跨多个产业的领头公司为支持IPv6都做出了重大投资，说明IPv6正从实验室中逐渐地应用到商业生产环境中。

## 4.1 制订一个适当的计划

IPv6的部署已经不是一个“如果”而是一个“何时”的问题了，那么如何做好准备呢？无非就是制订一个计划。本书的目的就是为你在计划过程中提供帮助。拥有一个可靠、经过深思熟虑的合适的计划，将会使部署过程流

水线化，并减少部署时出现意外情况。部署 IPv6 的计划过程现在就应该开始了。

没有两个网络是完全相同的，在为你的网络部署 IPv6 时，网络的具体需求和基础设施元素必须要得到妥善处理。因此，你组织内 IPv6 的部署不可能是一项简单的任务。另一方面，如果你保持更新不算太老的路由器、交换机和操作系统版本等网络基础设施，那你可能已经处在正确的道路上了。由于网络内 TCP/IP 的普遍性质，通过网络核心的基础，IP 如同织物一样被编织进去，所有受影响的群体内的业务也都必须参与进去。制定的计划必须包括所有主要的 IT 部门，包括基础设施、网络、安全和应用程序。对于任何参与管理、监控和采购（如虚拟主机、外部 DNS）等 IT 服务的 IT 合作伙伴，当这些组织随着 IPv6 的部署而功能可能受影响时，也需要它们参与进来。

作为计划的一部分，需要考虑的一些关键点如下：

- 开发一个业务案例，以概述 IPv6 的好处。要清楚这个案例是什么，为什么需要这个案例，以及与过渡有关的风险是什么。将此呈现给高级管理人员，并确保他们有一个坚定的认识从而愿意支持项目。开始时可以将焦点集中在你的网络的一个特定范围或者子集中，以“试验”该部署并获得一个成功的部署经历，从而为进一步的部署起到重要的影响，这是非常有意义的。这个过程在本书第 1 章中已经描述过了。
- 在你的组织内建立一个专门为 IPv6 负责的项目工作小组。要求成员来自于各个领域的 IT 组织及在你整个组织内的其他受影响部门。
- 了解你当前的网络环境。正如将在本章讨论的，执行你的网络发现系统或者使用你当前的 IP 地址管理（IP Address Management，IPAM）解决方案，来绘制出你现有的 IP 地址空间。盘点你现有的基础设施和应用。别忘了，在开始 IPv4/IPv6 共存之旅之前，需要知道你的 IPv4 现在到哪个地步了。
- 如果需要的话，借助外部资源。如果需要，考虑使用咨询服务，以帮助进行网络发现、IPv6 规划、项目管理和培训。
- 借助 IT 组织内关键的刷新间隔。在大部分组织内，计划中的正常的硬件和软件更新都可以作为你的优势，用已存在的 IT 更新计划来调整 IPv6 “更新”。
- 制订与 IPv6 相关的寻址、安全和网络管理的计划及路线图。在计划过程中，要尽早启动这个部分，因为这是很耗时的，而且必须要有组织内已经存在的安全架构的支持。接下来的三章将分别阐述这些领域。

## 4.2 IP 网络库存

### 4.2.1 IPv6 准备情况

使用 IPv6 准备情况来表示网络设备、操作系统、终端用户设备、应用程序或网络提供商的兼容性。要知道，这个概念不仅适用于硬件和软件，同时也包括了软件、过程和可操作性知识在内的整个 IT 功能。

你的组织应该建立一个具体到网络的 IPv6 准备程序。它应该包括下面几点：

• 对你当前 IP 地址库存和使用的回顾。

• 对你的网络和计算基础设施关于 IPv6 兼容性的评估。这应包括包含网络服务在内的硬件和软件。

• 对你的业务应用程序的审查，以确保它们已为 IPv6 做好准备，并且能够在 IPv6 网络上正确运行。

• 对你组织的 IPv6 技术能力的评估，涉及组织中包括工程、运营和支撑在内的所有角色。

• 对当前 IT 流程的回顾，确保它们包括 IPv6。

• 对安全策略、体系架构和实践的分析。

• 对通过网络和你的业务打交道的客户、合作伙伴和供应商系统的 IPv6 准备情况的测定。

对上述的每个评估范畴，你应该确认并跟踪其详细的标准，然后将其归类为以下三类：

• 已经支持 IPv6。

• 能够升级 IPv6。

• 不支持或升级 IPv6。

为了为你的网络协助这些评估领域，本书创建了一个 IPv6 准备情况模板。你可以使用此模板进行确认和分类 IPv6 准备情况。模板部分将在后面的章节中进行描述，完整的样本见本书附录；电子版本在网址 www. ipamworldwide. com 上可下载。

### 4.2.2 发现

首先是要确定你网络中的每一个组成部分。如果你已经有具体的网络，并计算库存信息，那么恭喜你！如果你没有或者你想核实这份库存信息，那么可以使用网络发现程序，其目标就是直接从每个设备中收集系统的详细信息。通过使用一个或多个如今许多可用的网络元素/设备发现工具进行你的网络设备的

发现，会使为 IPv6 就绪做准备的工作变得非常地简单，如以下几个工具：

- Netformx Discovery™[70]
- HP DDMI（Discovery and Dependency Mapping Inventory）[71]
- OPNET NetMapper®[72]

这些工具和类似的工具将为你提供网络中每个元素的重要信息，包括供应商、硬件版本及运行在网络设备上的操作系统/IOS 版本。

### 4.2.3 IPv6 评估

接下来的任务，就是将你的网络库存，包括网络和计算机设备、软件和应用程序的型号/版本列出清单，与已发布的每一项 IPv6 功能进行比较。由于需要把计算机设备和软件与供应商的数据表和 IPv6 准备情况报表进行比较，因此就算有的话也很少有工具可以自动完成这个手动和烦琐的过程。

随着评估阶段分析你当前网络的硬件、软件和应用程序的相关工作的进行，你可以借助一些资源，它们可以帮助你确定和评估你的设备和应用的 IPv6 功能。这些资源包括供应商信息、互操作性测试实验结果，以及普通的关于 IPv6 准备情况的信息网站。这里列出了一些可以帮助你的资源：

- IPv6 应用程序兼容性列表——美国威斯康星大学麦迪逊分校（http：//kb. wisc. edu/helpdesk//page. php？id = 11691）。一个关于 IPv6 软件应用程序兼容性版本信息和供应商报表的集合。

- IPv6 应用支持的比较——维基百科（http：//en. wikipedia. org/wiki/Comparison_of_IPv6_application_support）。一个关于 IPv6 软件应用程序兼容性版本信息的列表。信息是有些过时的，不过网站确实有到应用程序开发人员网站的链接，这些网站有更多当前信息。

- IPv6 标准——IETF 的 IPv6 和 IPv6 维护组（http：//www. ipv6- to- standard. org/）。一个由 IETF 维护的数据库，包含一个启用了 IPv6 且取得了成功运作 IPv6 经验产品的列表。

- IPv6 论坛——“IPv6 Ready Logo”项目（www. ipv6ready. org）。提供一份“IPv6 Ready Logo”项目认可的软件和硬件设备列表。这是一份已经通过标识规范的测试并实现 IPv6 就绪状态的软件和硬件列表。

- 设备和软件供应商——大多数主要的设备和软件供应商都更新了它们与 IPv6 准备情况相关的网站和 IPv6 相关的符合性声明。与供应商紧密合作，以确保如果它们还没完全与 IPv6 兼容，你也能了解它们的 IPv6 发展路线图。我们需要真正的 IPv6 支持，并仔细了解部分兼容的全部含义。

如果你愿意的话，不同的供应商会为你提供 IPv6 的评估服务来分担这项任务。IPv6 部署评估不仅包括能够运行 IPv6 的硬件和软件。评估必须包括你的 IT

结构的所有方面，包括以下各方面：

- IP 寻址。
- 关键的网络服务。
- 网络基础设备设施。
- 软件应用，包括现成的和定制的业务应用、运营支撑系统（Operational Support System，OSS）应用、监控系统、供应系统及其他支持系统。
- 技术技能和知识。
- IT 相关的过程，包括网络管理。
- 安全策略，体系架构和实践。
- 通过网络与你进行业务交互的客户、合作伙伴和供应商系统。

#### 4.2.3.1 IP 地址评估

在 IPv6 的评估过程中，第一步就是需要确定当前 IP 地址的分配。你必须完全了解已经部署的 IPv4 空间，如何分配和在何处分配，以及目前的使用率情况。如果你正使用免费的或者市售的 IPAM 工具来管理你的 IP 地址空间，那么对于收集这部分数据你应该可以有一个良好的开端。如果没有的话，你就需要进行 IP 地址发现，并收集这部分数据。你的 IP 地址基础设施的评估必须包括以下几点：

- 根块的分配。收集你的组织当前拥有的所有分配列表（IPv4 和 IPv6，如果有的话），并将这些信息放置到存储库。这不仅包括你从 RIR“租赁”的公共地址空间，还有私有空间，也就是 RFC 1918 空间。对存储这部分信息而言，IPAM 工具是一个完美的存储库，不过即使是一个简单的电子表格也是可以的。其中要包括被分配给这些块的每一块的技术和业务联系人以供将来参考。
- 为你每个根块逐条分配 IP 地址规划。此计划包括描述网络中 IP 地址空间如何分配的一份详细列表或图。通常情况下，IP 地址空间是根据应用程序和位置来进行分配的，而在大型网络内可能根据多个层次进行分层分配。这些“使用”的标识和每个 IP 地址分配上的分层，以及技术和管理联系人的信息，都是值得拥有的。
- 所有关于你 IP 地址分配的 IP 使用率信息。有了每个可用的 IP 地址分配的使用率信息，将会在设计阶段节省大量的时间。收集这些信息一般都通过使用网络扫描工具（捕捉静态地址分配）和通过从你已经存在的 DHCP 服务器中收集 DHCP 租赁信息来获取。目前，市场上有几个可用的工具和服务，使用这类工具和服务将有助于你的收集工作。
- 建立 IP 地址分配策略。如果你还没有在你的组织内使用标准策略，现在正是制定它们的时候。考虑使用模板或者与 IP 地址分配同样模式的方法，定义一个标准化的方法以使空间分配能够继续深入发展。层次分配对于维持有效的

路由是至关重要的。稍后将讨论相关的逻辑，并在本书第 5 章帮助指导你定义 IP 寻址策略。

图 4-1 所示为 IP 地址评估模板示例。使用该模板你能够列举出从每个 RIR 获取的根地址块。在块地址栏中用 CIDR 表示法列出每个块，并为每个块输入与各 RIR 状态相关的更新日期或结束日期。许多 RIR 有地址返还政策，所以随着时间的推移你可能要利用好这一点，尤其是你要更好地使用 IPv4 或者完全过渡到 IPv6（虽然这个最终状态需要好些年来完成）。如果你已经获得了一些 IPv6 地址，或者如果你正使用相当于 ULA 空间的私有地址进行实验，那么你也可以用根和层次分配来表示这些 IPv6 地址。

| 功能区 | 项目 | 块地址 | 使用/利用率 | 评估 | 下一步的计划 |
|---|---|---|---|---|---|
| IPv4地址空间 | | | | | |
| | ARIN根块分配 | | | | |
| | RIPE根块分配 | | | | |
| | APNIC根块分配 | | | | |
| | LACNIC根块分配 | | | | |
| | AfriNIC根块分配 | | | | |
| | RFC 1918根块分配 | | | | |
| IPv6地址空间 | | | | | |
| | ARIN根块分配 | | | | |
| | RIPE根块分配 | | | | |
| | APNIC根块分配 | | | | |
| | LACNIC根块分配 | | | | |
| | AfriNIC根块分配 | | | | |
| | ULA根块分配 | | | | |

图 4-1　IP 地址评估模板示例

“下一步的计划”这一列让你可以为一个项目放置一个检查标签，标签记录了当向 IPv6 迈进时需要的一些行为。例如，如果你还没有获得 IPv6 地址空间，就将 IPv6 地址空间区域中这项的需求表示在“下一步的计划”列中。图 4-2所示为 IP 地址根块评估示例。这个示例中给出了一对公共的 IPv4 块，一个 ARIN 的虚拟分配，一个 APNIC 的虚拟分配。另外，还给出了这些块当前的使用率，同时也给出了它们各自的更新日期及一些下一步的计划。在私有 IPv4 地址空间中，有一个 10.0.0.0/8 块和一对因公司收购而有的 192.168.0.0/16 块。对于每一项，还有其利用率和评估说明。为了 IPv6 部署工作，从 APNIC 中分配了一个 2001：db8：4af0：：/48 的 IPv6 块给这个组织。

为了补充你的 IP 地址评估，需要进一步的文档来描述你当前的 IP 地址规划，以图、列表或者电子表格的形式来描述一个公共的和私有的根块是如何在

| 功能区 | 项目 | 块地址 | 使用/利用率 | 评估 | 下一步的计划 |
|---|---|---|---|---|---|
| IPv4地址空间 | | | | | |
| | ARIN根块分配 | 192.0.2.0/24 | 62% | 3/17/2014过期 | 为IPv4因特网保留 |
| | RIPE根块分配 | | | | |
| | APNIC根块分配 | 198.51.100.0/24 | 99% | 10/22/2013过期 | 需要IPv6分配的补充 |
| | LACNIC根块分配 | | | | |
| | AfriNIC根块分配 | | | | |
| | RFC 1918根块分配 | 10.0.0.0/8 | 87% | 完全集成的 | |
| | | 192.168.0.0/16(主要的) | 72% | 分隔的网络 | |
| | | 192.168.0.0/16(收购的) | 69% | 被收购公司网络 | |
| IPv6地址空间 | | | | | |
| | ARIN根块分配 | | | | |
| | RIPE根块分配 | | | | |
| | APNIC根块分配 | 2001:db8:4af0::/48 | 1% | 初始分配 | 用于IPv6部署的块 |
| | LACNIC根块分配 | | | | |
| | AfriNIC根块分配 | | | | |
| | ULA根块分配 | | | | |

图 4-2 IP 地址根块评估示例

你的组织内进行分配的，就像下面 10.0.0.0/8 根块的示例一样。尽可能结合路由拓扑对你的地址空间建模。例如，如果实现了常见的三层核心分布访问拓扑，那就在每一层表示地址分配情况，就像图 4-3 所示的那样，通过缩进把每个地址块层次性地展现出来，最顶层的是全局分配（/8），下一层是大陆核心层的分配（/12），接下来是各自的地区分配（/16），然后是访问/子网的分配（/24）。

| 核心位置 | 地区 | 位置 | IPv4网络 |
|---|---|---|---|
| 全局分配 | | | 10.0.0.0/8 |
| 南美洲分配 | | | 10.0.0.0/12 |
| | 东部 | | 10.0.0.0/16 |
| | | 费城 | 10.0.0.0/24 |
| | | 蒙特利尔 | 10.0.1.0/24 |
| | | 华盛顿 | 10.0.2.0/24 |
| | 中部 | | 10.1.0.0/16 |
| | | 渥太华 | 10.1.0.0/24 |
| | | 休斯顿 | 10.1.0.1/24 |
| | | 丹佛 | 10.1.0.2/24 |
| | 西部 | | 10.2.0.0/16 |
| | | 旧金山 | 10.2.0.0/24 |
| | | 西雅图 | 10.2.0.1/24 |
| | | 圣地亚哥 | 10.2.0.2/24 |
| 欧洲分配 | | | 10.16.0.0/12 |
| | 东部 | | 10.16.0.0/16 |
| | | 柏林 | 10.16.0.0/24 |
| | | 基辅 | 10.16.1.0/24 |
| | 西部 | | 10.17.0.0/16 |
| | | 伦敦 | 10.17.0.0/24 |
| | | 巴黎 | 10.17.1.0/24 |
| | | 罗马 | 10.17.2.0/24 |

图 4-3 IP 地址层次分配示例

#### 4.2.3.2 网络基础设施模板

网络基础设施的模板提供了方便的跟踪表格，可用于记录和表示你的核心网络服务、路由和交换基础设施、终端用户设备、软件应用程序和有关系的客户或合作伙伴等 IPv6 准备情况。下面将依次讨论上述各个领域。图 4-4 所示为网络基础设施设估综合模板示例，可对每个网络组件进行标识（IP 地址、应用程序名称、序列号、MAC 地址，或者任何跟踪你的设备和应用程序的方式）。对于每个组件，你应该表示出提供该组件的供应商，以及硬件、操作系统和其本身“功能”的当前版本。例如，对于具有硬件平台和操作系统软件的服务器，给出安装在此服务器上的 DNS 版本。

评估需要定义每一个“子组件”的 IPv6 能力，以确定其功能是本来就完全支持 IPv6，还是需要升级或者根本不支持。通过检查评估模板上对应列中的项来找出此项答案。模板中也提供了表示任何 IPv6 限制或警告的内容，还有为 IPv6 升级或测试而需要的其他单元，以及过渡到具有 IPv6 能力的下一步计划。现在就来探讨这些领域的更多细节。

#### 4.2.3.3 网络服务

在这个升级过程的初期，你的关键网络服务也必须被评估。这些服务包括 DNS、DHCP、Radius 和基本网络所需的其他服务。如果这些基本的服务还未支持 IPv6 的话，它们必须升级到支持 IPv6。记住重要的一点，就是要注意当你在查询每一项服务的能力时，其中一些可能在为网络用户提供 IPv6 服务的同时也需要通过进行配置和管理使其支持 IPv4。也就是说，有些可能就功能方面上而言是双协议栈的。一般来说，不能将部署 IPv6 看做是替换 IPv4，应该预计将会在很长一段时间同时支持 IPv4 和 IPv6。对你的网络内的每个网络服务的评估必须要包括以下方面：

- 收集实现网络服务的当前供应商。如果网络服务架构还不是你整个网络计划或网络图的一部分的话，这是一个记录的好时机。
- 收集每个实现的网络服务的当前版本。指出每个版本是开源的或者是商业网络服务，并确保你了解网络服务的特定版本。
- 详细列出每个网络服务平台的当前硬件平台和操作系统。理解和记录这点很重要，因为尽管网络服务可能完全支持 IPv6，如果底层的硬件和操作系统不支持，那么你就不能够利用这些 IPv6 功能。
- 记录每个网络服务当前版本的 IPv6 功能和局限性。与每个网络服务供应商合作，或者收集开源网络服务有关的信息，从而记录当前版本的 IPv6 能力和局限性。如果需要的话，供应商应该能够为你提供一份 IPv6 准备情况说明。

准备情况模板的网络服务部分，提供了关于一些网络服务需要考虑的具体问题和要点信息项的关键字段。请一定不仅要注意关于 IPv6 传输的 IPv6 能力，

| 功能区 | 项目 | 项目ID | 当前厂商 | 当前版本 | 当前硬件与操作系统 | 操作系统支持IPv6? | 功能支持IPv6? | 评估：选择一项 | | | IPv6的具体限制 | 需要的其他单元 | 具有IPv6功能的下一步的计划 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 项目完全支持IPv6 | 项目支持IPv6升级 | 项目不支持IPv6 | | | |
| **网络服务** | | | | | | | | | | | | | |
| | DHCP | | | | | | | | | | | | |
| | DNS | | | | | | | | | | | | |
| | NTP/SNTP | | | | | | | | | | | | |
| | Radius/Diameter | | | | | | | | | | | | |
| | FTP | | | | | | | | | | | | |
| | TFTP | | | | | | | | | | | | |
| | Rsync | | | | | | | | | | | | |
| | SMTP/POP/IMAP | | | | | | | | | | | | |
| | HTTP | | | | | | | | | | | | |
| | 其他 | | | | | | | | | | | | |
| **网络基础设施** | | | | | | | | | | | | | |
| | 路由器 | | | | | | | | | | | | |
| | 核心交换机 | | | | | | | | | | | | |
| | 分布式/边缘交换机 | | | | | | | | | | | | |
| | 负载均衡器 | | | | | | | | | | | | |
| | 应用服务器 | | | | | | | | | | | | |
| | 防火墙 | | | | | | | | | | | | |
| | SAN/NAS存储系统 | | | | | | | | | | | | |
| | 无线接入点 | | | | | | | | | | | | |
| | IP电话服务器 | | | | | | | | | | | | |
| | 其他 | | | | | | | | | | | | |
| **终端用户/端点系统** | | | | | | | | | | | | | |
| | 台式机/笔记本/工作站 | | | | | | | | | | | | |
| | 平板电脑/PDA | | | | | | | | | | | | |
| | 智能手机 | | | | | | | | | | | | |
| | 其他手持设备 | | | | | | | | | | | | |
| | 打印机 | | | | | | | | | | | | |
| | 销售点设备 | | | | | | | | | | | | |
| | CPE设备 | | | | | | | | | | | | |
| | 其他 | | | | | | | | | | | | |
| **软件应用程序** | | | | | | | | | | | | | |
| | 商业应用1 | | | | | | | | | | | | |
| | 网络管理应用1 | | | | | | | | | | | | |
| | OSS应用1 | | | | | | | | | | | | |
| **客户/合作伙伴系统/链接** | | | | | | | | | | | | | |
| | 合作伙伴系统1 | | | | | | | | | | | | |
| | 合作伙伴系统2 | | | | | | | | | | | | |

图 4-4　网络基础设施评估综合模板示例

而且也要注意适当地在“应用层”支持 IPv6 地址。你可以在模板中重复多行，从而为每个网络服务提供复查其已服务的范围的机会，以帮助确定潜在的能力问题。

如果你每项服务使用多个供应商或者可能只有一个供应商，或如果你想要跟踪这种详细程度，需要的话对于一个给定的网络服务你可以有多个行。评估列应该被用来表示每个网络服务是否能够同时满足 IPv4 和 IPv6 的要求，可通过软件升级或硬件升级，以及软件和硬件同时升级；或者选择不支持。

你可以检查一个评估列，然后将这样的规定表示在“下一步的计划”列，并适当地在采购清单或者“需要的其他单元”列中增加一项。图 4-5 所示为包含一组 ISC 和微软 DHCP 服务器的网络评估示例。在此，我们“放大”了工作表的一部分，其中记录了每个服务的相关信息、各自的 IPv6 能力和评估，以及下一步的计划。

在某些情况下，即使当前的网络服务可同时支持 IPv4 和 IPv6，你可能还是想要添加一项网络服务到采购清单中。这一策略旨在将修改一个工作的 IPv4 配置，以使其同时支持未经证实的 IPv4 和 IPv6 配置的风险降到最低。你的评估应该说明这一点，不管添加一个新的网络服务是由于功能上的原因，还是需要/想要降低风险。隔离 IPv6 网络服务流量可以帮助其专注在 IPv6 操作中而不影响 IPv4 操作，直到达到一个较合适的 IPv6 水平。另外，如果你没有可用的服务器用于部署的测试阶段，那你可能非常渴望为你的测试实验室采购一个或者多个网络服务服务器。总之，如果一个特定的网络服务需要任何形式的升级、更换、补充或实验室安装，那就在“下一步的计划”列中记录此项内容。

#### 4.2.3.4　网络基础设施设备

虽然许多硬件和应用程序供应商在各种各样的功能上支持 IPv6，但是可能需要硬件或软件升级。大部分不算太老式的网络基础设备和软件已经至少在一定程度上支持 IPv6 了。要确定打算使用的特定 IPv6 功能，如移动 IPv6 或地址自动配置，并核实每个供应商提供的功能支持。

供应商应该能够为你提供所需要的 IPv6 准备情况和它们设备的功能的所有细节。也有一些在因特网上提供的可用资源，特别是 IPv6 的互操作性实验室已经证明了 IPv6 测试和互操作性测试的结果。为了完成你的评估，请完成以下几点：

- 详细列出实现网络设备当前的供应商和型号。如果你的组织还没有适当的网络设备资产系统，这可能是去实现一个的好时机。
- 列举出在网络设备上运行的操作系统和软件的当前版本。如果你的组织尚未标准化操作系统代码等级，这可能也是个好时机。
- 确定当前设备的 IPv6 能力和局限性。与供应商合作，或通过收集开源网

| 项目 | 项目ID | 当前厂商 | 当前版本 | 当前硬件与操作系统 | 操作系统支持IPv6? | 功能支持IPv6? | 评估：选择一项 | | | IPv6的具体限制 | 需要的其他单元 | 具有IPv6功能的下一步的计划 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 项目完全支持IPv6 | 项目支持IPv6升级 | 项目不支持IPv6 | | | |
| **服务** | | | | | | | | | | | | |
| DHCP | 10.200.0.11 | ISC | 3.2 | RHEL v5 | 是 | 否 | | √ | | | | 升级到ISC 4.2 |
| | 10.16.35.98 | ISC | 4.1 | RHEL v5 | 是 | 是 | √ | | | DHCPv4或v6，不同时 | 1 | |
| | 172.19.23.55 | 微软 | 2003 | Windows 2003 | 是 | 否 | | √ | | | | 升级到2008R2 |
| | 10.104.39.213 | 微软 | 2008R2 | Windows 2008R2 | 是 | 是 | √ | | | | 2 | |

图 4-5　包含一组 ISC 和微软 DHCP 服务器的网络评估示例

络服务有关的信息，记录当前版本的 IPv6 能力和局限。如果需要的话供应商应该能够为你提供一份 IPv6 准备情况的说明。

正如模板中的网络服务部分一样，可能需要为你的网络中实现的每个网络基础设施重复多行以单独跟踪每个设备和应用，或者你可能想通过类型和版本进行简单的概括。指出传输层和处理层的 IPv6 能力。对于那些需要软件和/或硬件升级或替换的元素，在“下一步的计划”列中指出具体细节。同时在“需要的其他单元”列中指出任何用于备用或实验室安装的所需采购信息。

#### 4.2.3.5 终端用户/端点设备

大部分组织内的大多数设备都是由终端用户和端点设备组成的。终端用户设备还常包括最广泛多样性的设备类型，特别是如今公司网络中自带设备的激增，其数量和多样性也在增长。从服务供应商的角度来看，端点设备可能包括客户端设备，如客户边缘路由器，或无线电、光纤、DSL，或电缆调制解调器。

与所有连接到你的网络的其他硬件和软件一样，所有终端用户和端点设备应该在可能的范围内被评估。不像网络设备，大多数组织在其上都有标准化的首选供应商和型号，端点设备可能变化相当多样，而且数量庞大的不同设备类型可能需要一些额外工作以进行适当的 IPv6 准备情况评估。我们建议首先要关注一些关键的设备，如笔记本、台式机、打印机、VoIP 电话、销售终端及手持设备等对业务非常关键的设备。

为了完成你的评估，请完成以下几点：

- 详细列出正在你的网络上使用的终端用户/端点设备的当前供应商和型号。如果为员工提供笔记本和台式机，那么你很可能已经有一个可以查询的资产数据库了。
- 列举出在这些设备上运行的当前操作系统和软件版本。首先要集中在最关键的系统上。
- 确定当前设备的 IPv6 能力和局限性。与供应商合作，或通过收集开源网络服务有关的信息，记录当前版本的 IPv6 能力和局限性。如果需要的话供应商应该能够为你提供一份 IPv6 准备情况的说明。

通过逐行完成评估表格中的项，对每个终端用户设备类型的版本进行评估，并完成每一项评估。测试每个设备类型 IPv4 和 IPv6 的协议栈，以及计划中的隧道技术的支持度。如果打算使用翻译技术，那么你的预部署测试计划中要包括对常见的终端用户设备的测试，以确保实现正常的通信，这将在本书第 8 章讨论。

为每一设备类型适当地测试其通用或共同的应用，以识别出任何 IPv6 地址表示或 DNS 处理的问题。对于识别出的任何问题，咨询相应的供应商。对于任何检测到的缺陷，将其记录在相应“具体的 IPv6 限制”列中并将解决途径记录

在“下一步的计划”列中。此外，可能需要采购一些能满足你的IPv6准备情况评估的设备类型，不过要提醒你这些设备类型需要预部署测试。

#### 4.2.3.6 软件应用

正如大多数人非常清楚的，将关键业务应用、管理软件或自定义编码系统升级到一个新的版本，本身就可以成为一项重大工程。在整个过程的早期就获得一份准确且完整的软件应用清单，是非常重要的。由于需要等待供应商完成发布版本，并在内部进行测试，然后才方便项目的部署，故而升级软件的交付周期可能会很长。幸运的是，IPv6已经提上日程很长一段时间了，大多数供应商已经添加了或者在这个过程中添加了IPv6支持。

评估需要包括以下方面：

- 详细列出实现的软件的当前供应商。如果你的组织还没有适当的软件资产系统，这可能是去实现一个的好时机。
- 表示出每个处于运行的软件的当前版本。
- 记录下各组件的依赖关系，如数据库供应商和版本，并评估每个组件的IPv6准备情况。
- 列出每个应用的功能。
- 通过与供应商或软件开发人员合作，识别并记录当前软件的IPv6能力和局限性。如果需要的话，供应商应该能够为你提供一份IPv6准备情况的说明。

与网络服务和网络基础设施一样，可以逐条检查网络中使用的每个业务、网络管理和OSS应用。对于每一项，确定其IPv6能力和局限性，并确定为实现对IPv6的支持所需采取的行动。确保测试了4.2.3.5节提到的终端用户设备中应用程序的“客户”端。别忘了为需要升级、补充或更换的应用更新“下一步的计划”列。

#### 4.2.3.7 客户、合作伙伴和供应商系统

任何通过网络与你的业务交互的客户、合作伙伴和供应商的系统，都需要检查其兼容性。在许多情况下，需要与你的合作伙伴和供应商制定一个计划，以允许IPv6流量通过。确定为使合作伙伴和供应商系统符合IPv6规定而需要的具体升级和/或配置参数的变化，并将那些需要采取的行动记录在“下一步的计划”列中。

#### 4.2.3.8 过程、实践与工作人员准备情况

除了IP地址和网络基础设施之外，第三个主要的评估领域就是需要评估现有的管理和安全实践与过程，同时还有IPv6工作人员的整体准备情况。图4-6所示为过程与工作人员准备情况评估模板示例。

在对应条目中详细列出每个过程和工作人员。对于每个过程，确定IPv6是否被考虑进去了。IPv6需要被以某种形式添加到大部分的安全和管理过程与实践中。根据各自的角色，工作人员如果不是IPv6方面的专家，那么也要在其他对应的方面非常精通。

| 功能域 | 项目 | IPv6就绪或认证? | 混合IPv6下一步的计划 |
|---|---|---|---|
| IT管理过程 | | | |
| | 过程1 | | |
| | 过程2 | | |
| | | | |
| 安全过程 | | | |
| | 过程1 | | |
| | 过程2 | | |
| | | | |
| 工作人员准备情况 | | | |
| | 网络架构师1 | | |
| | 网络工程师/分析师1 | | |
| | 网络技术员1 | | |
| | IT服务台工作人员1 | | |

图4-6 过程与工作人员准备情况评估模板示例

#### 4.2.3.9 技术技能与知识

工程师培训是IPv6部署的一个重要组成部分。虽然也有一些用于IPv6技术培训的可用资源，但是IPv6论坛发起了一项IPv6教育认证计划，以促进IPv6培训计划的一致性。该计划的目的是帮助鼓励和加快IPv6的教育，并促进加快IPv6的采纳速度。有几家厂商提供了IPv6的课程。要认准具有IPv6培训和获得IPv6认证标志的课程的供应商。IPv6论坛提供了一个包含IPv6认证讲师名单的数据库。值得注意的一些组织如下：

- IPv6college——Tonex的一个业务部门。
- Nephos6。
- 6Deploy。

对于每个位置类型，列出每一个工作人员及他/她的目前IPv6的专业知识水平，还有达到所期望的水平所需要的相应培训。如果有工作人员已经达到了他们位置上所需的IPv6专业知识水平，那么你就可以不填“下一步的计划”列，否则在该栏里注明推荐的IPv6培训课程。

#### 4.2.3.10 IT相关的过程

每个IT组织都有IT相关的过程，有些比其他组织更规范。如果使用ITIL程序来管理IT网络，那么你就可以列举出每个过程域，并评估其IPv6的准备情况。确保你的计划阶段要包括这些过程，并进行修正使其包含IPv6相关的支持。在“下一步的计划”中，要对每个过程定义需要更新的文档，以及与受影响团队所需的通信沟通。

#### 4.2.3.11 安全策略、体系架构和实践

虽然IPv6是将安全性考虑在内进行开发的，但把IPv6引入到网络中仍将对网络现有的安全性提出重大的挑战。你组织的安全架构需要进行审查和修改以

使其包括 IPv6。本书第 6 章将对安全系统配置方面详细地讨论安全性。而此时，先确认能够处理 IPv6 数据报。当读本书第 6 章时，你可能会发现需要详细的过滤功能，以推动进一步细化对升级或替换安全硬件与/或软件的评估。审查并评估以下方面：

- IPv4 部署架构问题。审查与你的部署类型有关的具体漏洞，这将作为一个重要的安全基线。
- IPv6 入侵检测系统。
- IPv6 防火墙能力。
- 专门针对 IPv6 安全的软件/硬件补丁。

图 4-7 所示为已部分完成的过程和工作人员评估示例。

| 功能域 | 项目 | IPv6就绪或认证? | 混合IPv6下一步的计划 |
|---|---|---|---|
| IT管理过程 | | | |
| | 发布管理 | 是 | 过程文档已更新；需要实验验证 |
| | 配置管理 | 否 | IPv6 配置方面需要被加到这一过程，并测试 |
| | | | |
| 安全过程 | | | |
| | 监控防火墙日志 | 更新了IPv6策略 | 需要运行测试用例来验证防火墙日志并记录方法和步骤 |
| | 锁定过程 | 更新了IPv6策略 | 在部署之前需要在实验室进行测试 |
| | | | |
| 工作人员准备情况 | | | |
| | Steve Jones | 是 | 2010年获得认证—可参加一个进修班 |
| | Mary Thompson | 是 | 2012年获得认证 |
| | Julia Starkly | 否 | 安排培训—概述，网络工程 |
| | Greg Libar | 否 | 安排培训—概述，网络工程 |
| | Mark Alexander | 否 | 安排培训—概述，帮助台 |

图 4-7 已部分完成的过程与工作人员评估示例

## 4.3 IPv6 待办事件清单

一旦完成了评估阶段，你应该有一份网络中所有硬件、软件、应用和终端用户设备的统一详细说明。在每份完成的评估模板中，“下一步的计划”列提供了一份关于所有需要跨不同部门的下一步行动内容的摘要。基于对每个这些项目的评估，为每个需要进行升级、更换、补充或实验室采购，从而实现遵守 IPv6 和/或支持 IPv6 部署测试的项目。你可以创建一个 IPv6 的补救或“待办事项”列表。过程更新和培训的下一步的计划也是待办事件列表的关键部分。

对于那些需要采购的待办项目，你可以生成一个需要定价的采购清单。定价需要在一个较高的水平上，以确定潜在的资本成本。随着收集到了更多的细

节，该信息应该被反馈到实际的业务案例中，从而进一步优化完善。由于高成本或缺乏利益的缘故，你可能会决定放弃某些非关键系统。这种“非 IPv6 兼容的元素”应当记录在你的网络资源库中。为了保持在预算内，你可能还需要相应缩减所需的实验室采购开销。理想情况下，你应该要优先考虑采购清单，从而在必要时帮助确定从最好的情况到必要的最低投入资金的范围。你的采购清单总额可以呈现给组织内的关键利益相关者以获取批准。不要忘记利用上你 IT 组织内的计算更新时间间隔。发生在大多数组织内的正常的 IT 更新可以被充分的利用，并与已有的整个 IT 计划保持一致。

## 4.4 IPv6 准备情况评估总结

图 4-8 所示的 IPv6 准备情况评估基本过程，以流程图的形式进行了总结。清查你的 IP 网络和包括过程与人在内的支撑结构的方方面面，对于定义你的出发点及在你的网络和/或你所选择的网络范围的内部署 IPv6 的所做的事项来说，是非常关键的。

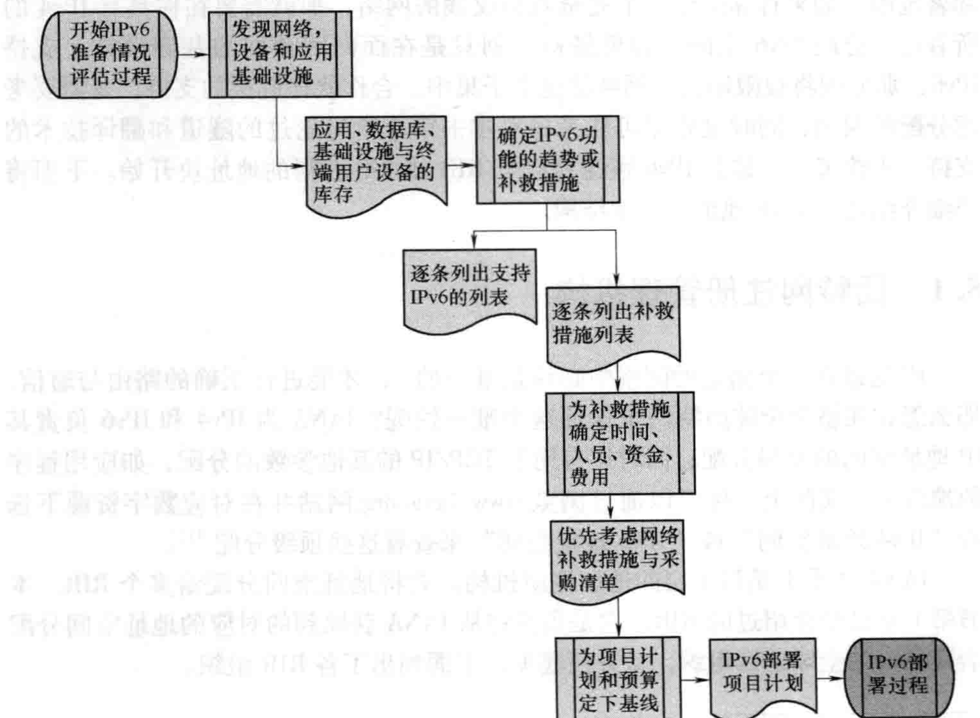



图 4-8　IPv6 准备情况评估基本过程

# 第5章 IPv6地址规划

IPv6准备情况评估阶段输出的是一个完整的网络库存清单，列出所有的基础设施和用户设备，以及IPv4和IPv6地址分配的库存。对应网络中的地址块（聚合），子网和DHCP地址池，该IP地址库存清单应该可以被拓扑映射到IP网络中。许多IP地址管理（IPAM）解决方案集成了这个功能，所以发现过程仅是验证IPAM数据库。如果你还未利用IPAM解决方案，或你的IPAM解决方案从地址块到子网再到地址池或单个地址的分配都不覆盖完整的IP地址的生命周期，那么这个映射将需要手动进行。如果地址块和子网的分配已记录在电子表格或其他存储库中，则这项信息将是此过程很有用的输入。

首先需要一个自顶向下的IP地址规划作为IPv4"计划"的基线，从而为你的IPv6地址的覆盖过程提供坚实的基础。你的IPv6覆盖程度将取决于所选择的部署范围。如果打算操作一个完全双协议栈的网络，那就需要在你具有IPv4的所有地方分配IPv6空间。如果最初计划只是在面向因特网的基础设施上支持IPv6，那分配将被限制在你网络的这个子集中。合作伙伴链接的支持，是需要考虑分配的问题，同时也要尽可能考虑对本书第3章讨论过的隧道和翻译技术的支持。不管怎样，你的IPv6覆盖将以从RIR或ISP获得的地址块开始。下面将详细介绍这个IPv6地址的层次结构。

## 5.1 因特网注册管理机构

IP地址在一个给定的网络中必须是唯一的⊖，才能进行正确的路由与通信。那么怎么在整个全球因特网中保证这个唯一性呢？IANA为IPv4和IPv6负责其IP地址空间的全局分配；同时还有用于TCP/IP的其他参数的分配，如应用程序的端口号。实际上，你可以通过浏览www.iana.org网站并在对应数字资源下选择"IPv4地址空间"或"IPv6地址空间"来查看这些顶级分配[73]。

IANA本质上是最上层的地址注册机构，它将地址空间分配给多个RIR。本书第1章已经介绍过的RIR，它是负责将从IANA获取到的对应的地址空间分配给其各自的全球地区组织，为方便起见，下面列出了各RIR组织。

---

⊖ 这一陈述的一个例外是任播地址通常是分配给多个主机的，而多播地址同样也是共享的。本语句适用于单播地址。

• 非洲网络信息中心（African Network Information Centre，AfriNIC）。非洲地区[74]。

• 亚太网络信息中心（Asia Pacific Network Information Centre，APNIC）。亚洲/太平洋地区[75]。

• 美网因特网地址注册管理组织（American Registry for Internet Numbers，ARIN）。北美地区，包括波多黎各和部分加勒比地区[76]。

• 拉丁美洲和加勒比地区 IP 地址注册管理机构（Regional Latin American and Caribbean IP Address Registry，LACNIC）。拉丁美洲和加勒比群岛[77]。

• 欧洲 IP 网络资讯中心（Reseaux IP Europeens Network Coordination Centre，RIPE NCC）。欧洲、中东和中亚[78]。

RIR 系统的目标如下：

• 唯一性。每个用于全球因特网路由的 IP 地址必须是全世界唯一的。

• 注册。一个可公开访问的 IP 地址分配的注册表，能够消除歧义，并在解决纷争时提供帮助。这个注册表被称为查询数据库。如今已有很多的查询数据库，不仅由 RIR 管理，同时也由 LIR/ISP 管理各自的地址空间。

• 聚合性。分层分配地址空间，以确保 IP 流量的正确路由。如果没有聚合性，路由表将会变得支离破碎，最终可能会在因特网中造成巨大瓶颈。聚合性被认为是 IPv6 分配最重要的目标。

• 避免浪费。特别是对于 IPv4，但对 IPv6 也一样，地址空间需要根据实际使用情况的要求进行分配。

• 公平性。在真实地址需求而非长期“计划”的基础上公平地进行地址分配。

• 简化管理开销。简化请求和获取初始分配与后续分配的过程。

每个组织要遵循这些目标做好它们自己的网络，特别是唯一性、注册（跟踪）、聚合性和避免浪费这几点。本章将讨论的分配方法，类似把地址块分配给 RIR，再通过 RIR 分配给国家或地方的互联网注册管理机构，然后再依次分配给服务提供商和终端用户。RIR 的分配指南记录见 RFC 2050（本书参考文献［4］）。一般的地址分配层次结构是自顶向下的 IP 地址分配，如图 5-1 所示。国家互联网注册管理机构类似地方互联网注册管理机构，但是在国家层面上组织的。

早在 20 世纪 80 年代和 90 年代初，许多企业（图 5-1 所示的终端用户）都是直接从 RIR 获得地址空间的。然而，在转变为 CIDR 寻址以提供进一步的地址分配责任下放的过程中，额外的 LIR/ISP 层被加入进来了。如今，大部分组织都是从 LIR 或 ISP 获取地址空间的。尽管 RIR 推荐使用一致性策略以最大限度地提高效率，获取这些地址空间的过程一般是由与你有业务往来的 LIR/ISP 决定的。

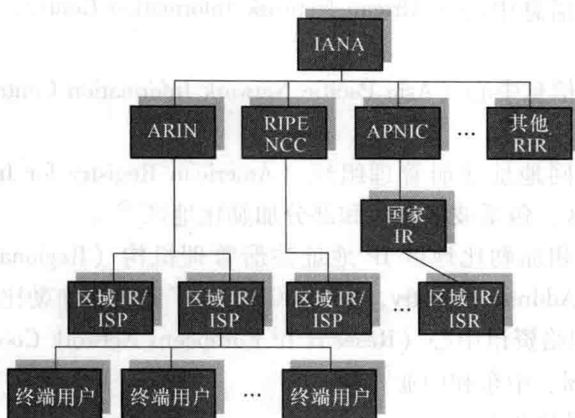



图 5-1 自顶向下的 IP 地址分配[4]

当空间被分配给一个 ISP 时，该 ISP 可能就会在因特网上公布其地址空间。加入的 LIR/ISP 层有助于在因特网上的路由聚合。由相同 ISP 服务的多个客户在因特网上可以归纳为一个路由。如果业务发展良好，且 LIR/ISP 需要更多的地址空间，则其可以从它们的 RIR 那里申请额外的空间。每个 RIR 一般都有为满足地址请求而自定义的过程，因此，请咨询你所在地区的 RIR 以获取更多细节。

## 5.1.1 RIR 地址分配策略

从 RIR 角度上看，RIR 分配地址空间给多个 LIR/ISP，然后 LIR/ISP 再把地址空间指派给它们的客户。从技术上讲，术语分配是指对于一个 IP 地址块，将其作为一个地址空间“池”，可以从这个地址空间“池”提取地址指派给客户。之后，客户就可以使用已分配的地址空间，从中分配块和子网，然后从已分配的子网中指派 IP 地址给单个主机。这种分配和指派机制是基于即将在本章描述的过程的，其与层次分配过程是一致的。然而，RIR 是将分配的空间与指派的空间区分开来的，因为指派的空间包含正在使用的地址；而分配的空间是用于进行分配的地址池，一开始为未使用，但理论上随着时间的推移和指派空间数量的增长，其利用率也会随着增长。从技术上讲，RIR 将分配的空间与指派的空间都看成是在使用中的，但保留了审计实际地址利用率能力，以便处理每个 LIR/ISP 后续的额外分配请求。

## 5.1.2 地址分配效率

在 IPv6 开发期间，许多研究都是集中在 128 位地址大小。虽然 IPv4 提供了一个 32 位的地址字段，从而提供了理论上最大的 $2^{32}$ 个地址或超过 42 亿个的地址，但实际上理论的最大值远远小于 42 亿。这是由于地址空间的分配是从网络

的多个层次，然后是子网，最后是主机进行分层分配的。RFC 1715（即本书参考文献［79］）提供了一种分析地址分配效率的方法，提出了将一个对数函数式作为分配效率的度量，将其定义为 H 比率：

$$H = \frac{\log_{10}(\text{对象的数目})}{\text{可用位的数目}}$$

根据网站“因特网世界统计”的数据[1]，如今有大概 24 亿的网络用户，因此现在的 H 比率为 0.293。42 亿 IP 地址的 100% 利用率对应的 H 比率为 0.301，因此现在的 H 比率相对来讲已经很高了。

IPv6 具有庞大的地址空间，其分配效率通过 HD 比率进行计算：

$$HD = \frac{\log_{10}(\text{分配对象的数目})}{\log_{10}(\text{可分配对象的最大数目})}$$

度量 IPv6 的 HD 比率公式中的“对象”指的是已分配的 IPv6 块地址(/48s)，这是用一个给定大小的 IPv6 前缀进行分配的。这些/48 地址块是 LIR/ISP 要分配给每个终端用户的。所以，如果一个具有/32 地址分配的 LIR/ISP 已经分配了 100 个/48 地址块，那么其 HD 比率为 $\log_{10}(100)/\log_{10}(65536) = 0.415$。

## 5.2　IPv6 地址规划

在一个组织内进行 IP 地址分配的主要目的是为每个网络内的终端节点提供 IP 地址，使得这些节点能够使用各种各样的媒体（数据、语音、视频等）与网络内的其他节点（也可能不是）及因特网节点（也可能不是）进行通信。除了考虑终端用户的寻址需求外，你的地址分配方案也必须考虑网络运行，以促进网络的管理和安全，这意味着是允许在内部进行通信或者是能够与因特网节点进行通信。

首先要检查第二支持者的需求，即网络运营团队的需求，考虑地址规划如何对路由器和防火墙的设置与策略产生影响。如果能简单地由被管设备的 IP 地址确定有关设备的信息，那么网络就更容易管理了。例如，在 IPv4 的世界里，大部分地址规划者都会为每个子网的路由器分配地址“.1”；如果使用更少地址相关的条目来定义诸如与访问控制列表（Access Control List，ACL）或路由处理（如语音与数据报处理）相关的策略，那么网络、路由和安全策略也会更容易管理。例如，一些组织定义了各种地址“类型”，以反映分配给基于应用程序的数据报处理的地址空间。通过充分的规划再加上可能的一些运气，IP 地址空间就会更容易管理，而且无需重新编码，经过多年后地址层次结构还能够保持完好，而重新编码的确是一件很痛苦的事。

用户群体主要为终端办公室或网络的“叶节点（leaf node）”所需的每种

类型的地址空间，推行整个网络拓扑的分配标准。在秉承独立性、聚合性和避免浪费的主要目标的同时，地址规划者必须为每一种地址类型分配足够的空间以满足容量需求。正如将要介绍的分配的例子，在制订地址分配计划和分层时，你不得不考虑取舍的问题。但首先要考虑各种不同的 IPv6 地址分配方法。

在部署 IPv6 的过程中，一个重要步骤就是从你的 RIR 或 ISP 请求一个 IPv6 地址块。虽然 IPv6 地址与 IPv4 地址的表示方式不同，但在网络中的分配过程本质上是相同的。主要的不同是 IPv4 是在十进制与二进制之间转换，而 IPv6 是在十六进制与二进制之间转换。通常用于 IPv4 分配的最小可用空闲块的最优分配过程，其实就是最佳分配算法。由于可用地址空间的巨大差异，IPv6 不仅支持一个类似的最佳分配算法，同时也支持一种稀疏的分配方法。后面将讨论这种稀疏的分配方法，以及一个可用来代替编号从 1 开始计数的简单子网编码的随机分配方法。下面会先谈谈这些分配方法，然后回到如何从实践的角度将这些方法应用到你的 IPv6 地址计划中。

## 5.3 IPv6 地址分配方法[⊖]

IPv6 和 IPv4 分配过程本质上是相同的，稍后会讨论一些捷径。最优的地址块分配方法需要最小可用空闲块的分配，这种方法被称为最佳分配算法。由于有庞大的可用 IPv6 地址空间，严格应用最佳分配算法可能是不必要的。

IETF 还定义了一个稀疏的 IPv6 地址分配方法，用于分配大小相等的块，而随着空间的增长，用随机分配方法来代替编号从 1 开始计数的简单子网编码。

下面将通过使用一个 IPv6 网络 2001:db8::/32 的例子来说明这些算法。实际上，一个/32（或任何一个）大小的全球单播分配需要一个区域因特网注册管理机构的资格预审，一个中等规模的企业组织不太可能会收到这样的分配。然而，例子将使用这个分配，以防止地址位数太长，以至于相关内容占用太多篇幅。稍后，下面将介绍一个更实际的/48 的分配例子。算法不管以/32 或是以/48的分配开始都是等效的，只是在/48 网络中将会有更多的前缀。

### 5.3.1 最佳分配方法

最佳分配方法力求用最小的可用地址块来分配所需的块大小。如果你使用的是大小相等的块分配（为简单起见建议如此，稍后将介绍），这种方法很简单。然而，为了最优地分配地址空间，分配的块应该不大于所需的大小。因此，

⊖ 这部分对 IPv6 分配的讨论是基于本书参考文献［81］的。

你可以在一次分配中分配一个/56 的块，然后在另一次分配中分配/61 的块。现在这种方法通常用于 IPv4 的块分配，因为存在节约地址空间的必要性。这个需求对具有巨大地址空间的 IPv6 而言是不迫切的，但尽管如此下面也会说明这个最佳分配方法过程。

最佳分配方法需要在你的地址层次结构中选择一个空闲的“父”块，使其能够完全满足所需的大小或者是更大尺寸的最小块。除非你能在你的大脑中进行十六进制的计算，不然这种方法通常需要在二进制域中进行分配。这是由于需要跟踪从相等或更大的块中分配可变大小的块。从稍后的例子中可以看到，这种处理的结果是除了分配所需的块外，也会有更多空闲的块，其本身可能会进一步被分配或被瓜分。通过网络 2001:0db8::/32 的例子，仔细看看这到底是如何发生的，以下以二进制的格式对其进行展开［部分地］：

**0010 0000 0000 0001 0000 1101 1011 1000** 0000 0000 0000 0000 0000…

现在假设需要从这个/32 的空间中分配 3 个/40 的网络。如果把/32 可用地址空间认为是一个可用来分配的饼，那么就可以对其进行分割。将饼对半切割从而产生两个/33 的块，列在第一个的是 2001:db8::/33，列在第二个的是 2001:db8:8000::/33。

**0010 0000 0000 0001 0000 1101 1011 1000 0**000 0000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 1**000 0000 0000 0000 0000…

原封不动地留下后半部分 2001:db8:8000::/33，以提供尽可能大的块用于后续的分配请求。同时，将 2001:db8::/33 部分分割成两个/34 的块，分别是 2001:db8::/34 和 2001:db8:4000::/34，即

**0010 0000 0000 0001 0000 1101 1011 1000 00**00 0000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 01**00 0000 0000 0000 0000…

现在原封不动地留下块 2001:db8:4000::/34，继续分割块 2001:db8::/34 为两个/35 的块，一直这样分割下去直到得到了一对/40 的块，过程如下：

**0010 0000 0000 0001 0000 1101 1011 1000 1**000 0000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 01**00 0000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 001**0 0000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0001** 0000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0000 1**000 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0000 01**00 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0000 001**0 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0001** 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0000** 0000 0000 0000…

这样，最佳分配方法依次减半地址空间以降到所需的大小。在本例子中，

现在有两个可用的/40 网络了（上面的阴影部分）。将其转换回十六进制就是 2001：db8：100：：/40 与 2001：db8：：/40。为了使用最佳分配方法分配所需的第三个/40 块，可以采用下一个最小的可用网络，在这个例子中就是 2001：db8：200：：/39，并将其分割为两个/40：

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0010** 0000 0000 0000…

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0011** 0000 0000 0000…

通过给下一个位赋值来将其进行对半分割，产生了两个/40 块。可以选择一个用于分配，而另一个将留下在以后分配。因此，分配的三个/40 块为 2001：db8：：/40，2001：db8：100：：/40 和 2001：db8：200：：/40。而另一个/40 的块，也就是 2001：db8：300：：/40，可用于以后的分配。图 5-2 所示从 1 个/32 网络中分割出 3 个/40 网络分配结果，图中用饼状图的形式说明了这个依次减半的过程。

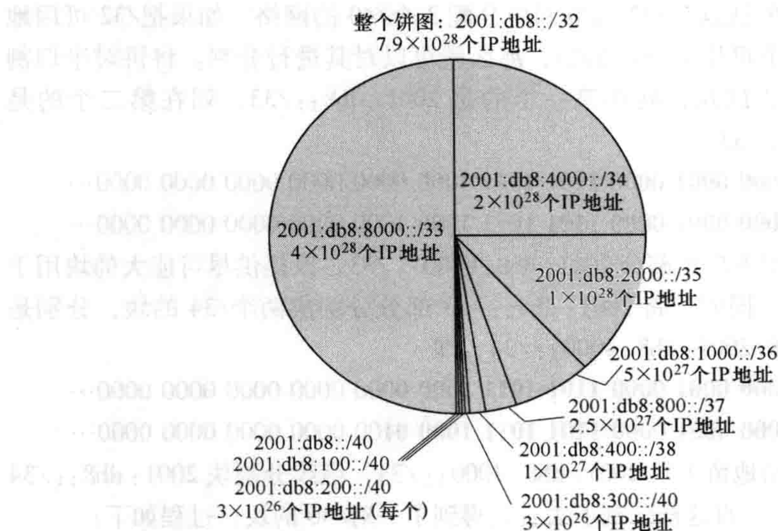



图 5-2 从 1 个/32 网络中分割出 3 个/40 网络的分配结果

当随后的分配请求中出现了一个/40、/38、/37、/36、/35、/34 或者/33 大小的块时，有随时可供分配的块来提供请求的地址容量，而无需分配多个可能不连续的块。对于一个其他大小的块的请求，如一个/44 的块，可以根据上面的分配过程以最小的空闲块 2001：db8：3000：：/40 开始进行分配。

当可变块大小根据策略进行分配时，最佳分配方法保留了较大的块。这种方法虽然算法复杂度较高，但其最好地利用了可用地址容量。因此对帮助服务提供商最大化地址利用率来说可能是很有意义的，但大多数企业可能会寻求其他下面将会描述的更易于管理的方法。

### 5.3.2　稀疏分配方法

在之前的从一个/32 块中分配一个/40 的块的算法中你会发现，网络的长度逐渐扩展到第 40 位；然后通过给网络的第 40 位赋值 0 或 1 作为一开始的两个/40网络。从本质上讲，这里依次处理了每个位，考虑将“1”作为空闲块而将“0”作为分配的块。然而，如果退后一步，将扩展/32 到/40 作为一个整体来考虑 8 个子网 ID 位，而不是递增地减半网络，实际上是通过对子网 ID 字段进行编号或计数来分配子网的，正如下面的阴影的加粗斜体位所示：

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0000** 0000 0000 0000…2001：db 8：：/40

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0001** 0000 0000 0000…2001：db 8：100：：/40

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0010** 0000 0000 0000…2001：db db 8：200：：/40

因此，如果你事先知道原始的/32 网络将仅被分割成均匀的/40 大小的块，那么就可以使用仅递增子网 ID 位的更简单的分配方法。下一个位/40 块分配的子网 ID 的值是 0000 0011、0000 0100、0000 0101，以此类推。

在某些网络中地址分配效率是最重要的，这种同样大小块的均匀性分配策略可能不适用，所以逐次减半的最佳分配方法可能会更合适。但从另一方面讲，稀疏分配方法提供了一种更简单的方法，并产生了虽非最优但却类似的好处。稀疏分配方法旨在使分配空间之间的地址空间最大化，以提供空间的增长。稀疏的分配方法也有着对半分配可用地址空间的作用，但不是持续这个过程直到下降到最小的大小，它需要在新的一半的边缘上分配下一个块。这样导致分配被散布开，而不是最优的分配。另外，其基本原理就是这种方法在丰富的 IPv6 空间中，通过在分配之间留下足够的空间以为分配的网络提供增长的空间。考虑这样一个例子，从 2001：db8：：/32 的空间中分配 3 个/40 块就会像这样：

**0010 0000 0000 0001 0000 1101 1011 1000 0000 0000** 0000 0000 0000…2001：db 8：：/40

**0010 0000 0000 0001 0000 1101 1011 1000 1000 0000** 0000 0000 0000…2001：db 8：8000：：/40

**0010 0000 0000 0001 0000 1101 1011 1000 0100 0000** 0000 0000 0000…2001：db db 8：4000：：/40

这分别转换为 2001：db8：：/40，2001：db8：8000：：/40 和 2001：db8：4000：：/40。这种分配可以使地址空间散布开，如图 5-3 所示。如果 2001：db8：8000：：/40 网络的接收者需要一个额外的分配，可以为其分配一个连续的或相

邻的块，即 2001：db8：8100：：/40。这个块在稀疏分配方法下将会是最后被分配的，所以它很有可能是可用的。在这种情况下，两个连续块的接收者就可以将它们的地址空间识别（或公布）为 2001：db8：8000：：/39。注意，这里的子网 ID 位是从左到右进行计数的，而不是用于“正常”传统计数从右到左的方法。

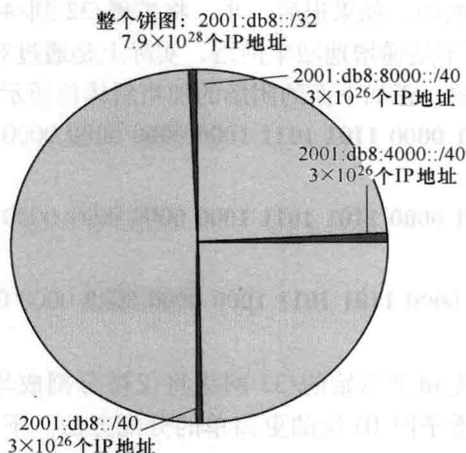



图 5-3 稀疏分配

RFC 3531（即本书参考文献［82］）描述了稀疏分配方法。由于网络分配期望遵循一个多层的分配层次，为了连续地分配，不同的实体可以使用多个连续网络位的集合。例如，一个因特网注册管理机构可能会分配第一个块给一个区域注册机构，该机构同样也会从该空间中分配块给一个服务提供商，服务提供商则可能会从该子空间中分配块给客户，而客户就可以在其网络中进一步分配。RFC 3531 建议较高级别的分配，如注册管理机构的分配，使用最左边开始计数或稀疏分配，最低级别的分配则使用最右边开始计数或最优分配，而其他中间级别的分配使用任一种皆可，或者甚至是一种最中心的分配方案。对于一个企业组织，可以使用稀疏分配方法来分配洲际或核心网络；在最顶层，不要过量使用地址空间，从而为未来的增长留出空间。

## 5.3.3 随机分配方法

随机分配方法随机选择一个在子网位数大小内的随机数来分配子网。使用之前从/32 中分配/40 的例子，这将会产生一个在 $0 \sim 2^8-1$（或 255）的随机数并对其进行分配，如果它仍可用的话。这种方法提供了一种横跨多个实体随机扩展分配的手段，其通常最适合于“相同大小”的分配。随机化的方法为不连续的块和以“1”开始的连续子网提供了一个“隐私”级别。要知道，随机分配

可能致使较大连续块的识别更加困难，而且会由于重新编号的原因而释放连续的空间。因此，在顶层更适宜进行稀疏分配，随机分配或最佳分配方法更适合在底部或子网分配级别的分配。

### 5.3.4 DHCPv6 前缀代理

DHCPv6 不仅可以用来给主机进行分配单个 IP 地址和/或相关的 IP 配置信息，也可以代理整个网络去请求路由设备。这种通过 DHCPv6 代理的形式被称为前缀代理。这种前缀代理的原始动机是因宽频服务提供商旨在用一种分层的方式来自动化授权 IPv6 子网（如/48 网络划分为/64 网络）给宽频用户的过程。一个在服务提供商面向用户的网络边缘的请求路由设备，会把一个通过 DHCPv6 协议发出的地址空间请求发布给一个代理路由器。

前缀代理过程使用的 DHCPv6 信息流，跟设备用来获得一个单独的 IPv6 地址所使用的基本的 DHCPv6 信息流相同。相应的 DHCPv6 消息中额外的信息可被用来确定一个被代理的合适网络。像 IPv6 地址一样，前缀具有优先和有效的使用期。请求路由器可以通过 DHCPv6 的 Renew 和 Rebind 信息来请求扩展使用期。请参考本书参考文献［11］，以获得更多有关 DHCPv6 的细节。

### 5.3.5 唯一本地地址空间

虽然 IPv6 并没有指定“私有”地址空间，但 ULA 空间本质上是等效的。通过使用 fc00::/7 前缀，设置 L 位为“1”（也就是 fd00::/8）来表明本地分配，再分配一个随机的 40 位全局 ID，你就可以得到一个/48 的前缀以供内网使用。就像 RFC 1918 中 IPv4 空间一样，ULA 编址的数据报，不能被路由到组织外，也就是因特网中。那么你是否应该分配 ULA 地址空间呢？对你的实验室实现来说为是；而通常对需要因特网访问的设备而言，则为否。这些设备将需要公共的 IPv6 地址分配和使用网络前缀转化（Network Prefix Translation，NPT），然而若将其作为实验的解决方案，则可能会降低其性能。

## 5.4 定义你自己的 IPv6 地址计划

既然定义了各种分配方法，那就讨论每一种方法可能会被用在哪些地方。下面将通过一个顶级核心网络，并假设其具有一个分层路由拓扑，来讨论分配问题。校园网或接入网络从顶级核心网络中分配地址，而本地子网则从校园网或接入网中进行分配。在所用的例子中，基于你网络大小和复杂性，这个简单的三层层次结构是可以被扩展为任何数量的层次的。第一个要考虑的是你是否要为网络中的每一个 IPv4 块或子网，映射一个 IPv6 块或子网。给定一个具有超

过 $18\times10^{18}$ 个地址的大型/64 IPv6 子网，肯定可以只使用一个/64 块就可以解决整个 IP 网络的编址问题。这当然会损害网络路由效率及其相关功能的优势，但某种形式的网络聚合可能会更加有意义。这并不是意味着要根据调试和解决路由和数据流问题而完全重新设计网络，网络的重新设计会在 IPv6 部署过程中引入复杂的额外层次。如果你的网络运行不佳，你可能需要先考虑重新设计和优化你的 IPv4 网络，然后再部署 IPv6。

假设你的 IPv4 网络虽不是最佳的但运行得还不错，那么你的 IPv4 地址可以作为你 IPv6 计划的一个可靠基础，这就是为什么定下你 IPv4 计划基线是如此重要。通常你的顶级 IPv6 分配将很可能会模拟 IPv4 的分配，为每个核心路由器分配一个聚合的地址块。核心路由器可能较少发布路由信息，因为每个相应的块在本地访问级别将会进一步被分配给下游。在分配 ISP 所分配的地址给核心路由器之前，一些网络管理员会首先根据应用程序的不同分割它们的空间，如 VoIP、数据、无线等。使用这种模板来分配地址有助于为特定应用程序通信配置网络路由策略和 ACL。例如，你可能会分割一个 ISP 提供的/48 块为 16 个/52 块来定义每个应用程序的空间。然后每个分配的/52 块可能会进一步被分配，假设每一个被分配为 16 个块，这样就为每个应用程序分配了 16 个/56 块。这种在你网络顶级的 IPv6 空间分配，应使用稀疏分配算法以防止路由表随着网络容量的增长而增长。

下面通过使用 ISP 分配 2001:db8:4af0::/48 的例子来说明这个过程。在应用层将迭代地应用稀疏分配，然后是核心区域层。在应用层，就像下面这样稀疏地分配第 13 个半字节或者 49 ~ 52 位。

| 核心分配 | 49 ~ 52 位 | 公共地址空间分配 |
|---|---|---|
| 数据 | 0 0 0 0 | 2001:db8:4af0::/52 |
| VoIP | 1 0 0 0 | 2001:db8:4af0:8000::/52 |
| 无线 | 0 1 0 0 | 2001:db8:4af0:4000::/52 |
| 管理 | 1 1 0 0 | 2001:db8:4af0:c000::/52 |

注意，你本来可以定义一个横跨整个网络的统一的各个应用策略清单。例如，如果源地址落在 2001:db8:4af0:8000::/52 上，就应用 VoIP 的数据报处理过程。由于有了稀疏分配算法，此时如果需要分配更多的“VoIP”地址空间，可以简单地分配一个 2001:db8:4af0:9000::/52（位 49-52 为 1001），它和 2001:db8:4af0:8000::/52 是相邻的。然后就可以仅从 2001:db8:4af0:8000::/52 到 2001:db8:4af0:8000::/51 更新我的路由、策略和 ACL。而下一步，为了分配这个 VoIP 空间到整个核心路由器，只需将接下来（第 14 个）的半个字节或 53 ~ 56 位分配给各大陆的地址空间：

| 子核心分配 | 53～56 位 | 公共地址空间分配 |
|---|---|---|
| 北美洲 | 0 0 0 0 | 2001:db8:4af0:8000::/52 |
| 欧洲 | 1 0 0 0 | 2001:db8:4af0:8800::/52 |
| 亚洲 | 0 1 0 0 | 2001:db8:4af0:8400::/52 |
| 南美洲 | 1 1 0 0 | 2001:db8:4af0:8c00::/52 |
| 非洲 | 0 0 1 0 | 2001:db8:4af0:8200::/52 |
| 澳洲 | 1 0 1 0 | 2001:db8:4af0:8a00::/52 |

注意，这种方法提供了一个网络地址到相应应用程序或地区的可视化映射。这是在例子中使用的半字节递增分配的关键优势。稍后，你就会发现一个具有 2001:db8:4af0:8400:: 地址分配的主机是一个位于亚洲地区的 VoIP 设备。因为第 13 个半字节的值为 8，表明其是 VoIP 应用；而第 14 个半字节的值为 4，表明其在亚洲。

对于一个/48 的分配，你实际上只有/48 到/64 之间 4 个半字节可以使用，所以要做相应的计划。你可以不使用半字节的方式，但十六进制的映射具有更大的挑战性，且你牺牲了可视化映射的好处。在本例中，对于剩下的 2 个半字节，可以分别为接入网络和本地网络各分配一个。在这些较低的分配级别中，根据你的安全考虑来决定顺序分配，或随机分配可能会更有意义。但在持续的结构化分配中，你使地址到类型和位置的可视化映射更容易了。对你所有的公共地址而言，ISP 前缀是一样的，所以根据子网 ID 部分来追踪，使这样映射更简化了。

对于之前已经讨论过的 2001:db8:4af0::/48 的 IPv6 分配，表 5-1 给出的示例是一个将其往下映射到/64 子网的例子。在这个例子中，在本书图 4-3 所示的 IPv4 地址评估结果的基础上，构建了自己的 IPv6 地址计划。注意，这里是如何使用不同的阴影来阐述每一个层次的，首先将全局的/48 块分配分割为应用层的/52 块分配（水平列表示），再下一层是核心的/56 块分配，然后是/60的地区分配，最后则是/64 的站点/子网。注意，可识别出地址 2001:db8:4af0:c812::/64 是在欧洲地区西部罗马的一个管理子网，因为“c812”在表 5-1的右下角每一项分别映射为管理（c）、欧洲（8）、欧洲西部（1），以及罗马（2）。

回到之前提到的子网合并，在分配的例子中，欧洲西部有 16 个可分配的子网，其中子网 2 在罗马中。如果在罗马办公室中已经有了 3 个 IPv4 子网，那么可以考虑将这些合并到一个 IPv6 子网中。此外，还要考虑路由限制，如果可以的话，还要考虑进行多子网分配的最初原因。可能是由于安全或者其他原因，保留多个子网是很有必要的。

表 5-1 IPv6 层次块分配示例

| 核心位置全局分配 | 地区 | 位置 | IPv4 网 10.0.0.0/8 | 数据网 2001:db8:4af0::/52 | VoIP 网 2001:db8:4af0:8000::/52 | 无线网 2001:db8:4af0:4000::/52 | 管理网 2001:db8:4af0:c000::/52 |
|---|---|---|---|---|---|---|---|
| 南美洲分配 | | | **10.0.0.0/12** | **2001:db8:4af0::/56** | **2001:db8:4af0:8000::/56** | **2001:db8:4af0:4000::/56** | **2001:db8:4af0:c000::/56** |
| | 东部 | | 10.0.0.0/16 | 2001:db8:4af0::/60 | 2001:db8:4af0:8000::/60 | 2001:db8:4af0:4000::/60 | 2001:db8:4af0:c000::/60 |
| | | 费城 | 10.0.0.0/24 | 2001:db8:4af0::/64 | 2001:db8:4af0:8000::/64 | 2001:db8:4af0:4000::/64 | 2001:db8:4af0:c000::/64 |
| | | 蒙特利尔 | 10.0.0.1/24 | 2001:db8:4af0:1::/64 | 2001:db8:4af0:8001::/64 | 2001:db8:4af0:4001::/64 | 2001:db8:4af0:c001::/64 |
| | | 华盛顿 | 10.0.0.2/24 | 2001:db8:4af0:2::/64 | 2001:db8:4af0:8002::/64 | 2001:db8:4af0:4002::/64 | 2001:db8:4af0:c002::/64 |
| | 中部 | | 10.1.0.0/16 | 2001:db8:4af0:10::/60 | 2001:db8:4af0:8010::/60 | 2001:db8:4af0:4010::/60 | 2001:db8:4af0:c010::/60 |
| | | 渥太华 | 10.1.0.0/24 | 2001:db8:4af0:10::/64 | 2001:db8:4af0:8010::/64 | 2001:db8:4af0:4010::/64 | 2001:db8:4af0:c010::/64 |
| | | 休斯顿 | 10.1.0.1/24 | 2001:db8:4af0:11::/64 | 2001:db8:4af0:8011::/64 | 2001:db8:4af0:4011::/64 | 2001:db8:4af0:c011::/64 |
| | | 丹佛 | 10.1.0.2/24 | 2001:db8:4af0:12::/64 | 2001:db8:4af0:8012::/64 | 2001:db8:4af0:4012::/64 | 2001:db8:4af0:c012::/64 |
| | 西部 | | 10.2.0.0/16 | 2001:db8:4af0:20::/60 | 2001:db8:4af0:8020::/60 | 2001:db8:4af0:4020::/60 | 2001:db8:4af0:c020::/60 |
| | | 旧金山 | 10.2.0.0/24 | 2001:db8:4af0:20::/64 | 2001:db8:4af0:8020::/64 | 2001:db8:4af0:4020::/64 | 2001:db8:4af0:c020::/64 |
| | | 西雅图 | 10.2.0.1/24 | 2001:db8:4af0:21::/64 | 2001:db8:4af0:8021::/64 | 2001:db8:4af0:4021::/64 | 2001:db8:4af0:c021::/64 |
| | | 圣地亚哥 | 10.2.0.2/24 | 2001:db8:4af0:22::/64 | 2001:db8:4af0:8022::/64 | 2001:db8:4af0:4022::/64 | 2001:db8:4af0:c022::/64 |
| 欧洲分配 | | | **10.16.0.0/12** | **2001:db8:4af0:800::/56** | **2001:db8:4af0:8800::/56** | **2001:db8:4af0:4800::/56** | **2001:db8:4af0:c800::/56** |
| | 东部 | | 10.16.0.0/16 | 2001:db8:4af0:800::/60 | 2001:db8:4af0:8800::/60 | 2001:db8:4af0:4800::/60 | 2001:db8:4af0:c800::/60 |
| | | 柏林 | 10.16.0.0/24 | 2001:db8:4af0:800::/64 | 2001:db8:4af0:8800::/64 | 2001:db8:4af0:4800::/64 | 2001:db8:4af0:c800::/64 |
| | | 基辅 | 10.16.1.0/24 | 2001:db8:4af0:801::/64 | 2001:db8:4af0:8801::/64 | 2001:db8:4af0:4801::/64 | 2001:db8:4af0:c801::/64 |
| | 西部 | | 10.17.0.0/16 | 2001:db8:4af0:810::/60 | 2001:db8:4af0:8810::/60 | 2001:db8:4af0:4810::/60 | 2001:db8:4af0:c810::/60 |
| | | 伦敦 | 10.17.0.0/24 | 2001:db8:4af0:810::/64 | 2001:db8:4af0:8810::/64 | 2001:db8:4af0:4810::/64 | 2001:db8:4af0:c810::/64 |
| | | 巴黎 | 10.17.1.0/24 | 2001:db8:4af0:811::/64 | 2001:db8:4af0:8811::/64 | 2001:db8:4af0:4811::/64 | 2001:db8:4af0:c811::/64 |
| | | 罗马 | 10.17.2.0/24 | 2001:db8:4af0:812::/64 | 2001:db8:4af0:8812::/64 | 2001:db8:4af0:4812::/64 | 2001:db8:4af0:c812::/64 |

DNS 是 IPv6 分配中另一个需要考虑的。如果你的 DNS 管理委托给了组织内的不同群体，通过一些事先的筹划，可简化相应反向区域的委托。在上面的例子中，如果你的 VoIP 团队管理着所有的 VoIP DNS，你可以将区域 8. 0. f. a. 4. 8. b. d. 0. 1. 0. 0. 2. ip6. arpa 委托给它。然而，如果欧洲的团队运行着跨应用的 DNS，使用上面的分配策略，你可能不得不委托到 16 个反向区域，也就是 8. x. 0. f. a. 4. 8. b. d. 0. 1. 0. 0. 2. ip6. arpa 中 x = {0 - f} 的区域。通常数据路由和策略超越了反向区域的考虑，但这是潜意识中要考虑的，同时这也帮助促进权衡分配策略与其在网络操作中的限制。

## 5.5　多重连接与 IP 地址空间

术语多重连接是指一个企业提供了多个（>1）到因特网的连接。图 5-4 所示的多重连接架构，是一个简单的架构。一个多重连接的策略提供了几个好处[84,85]：

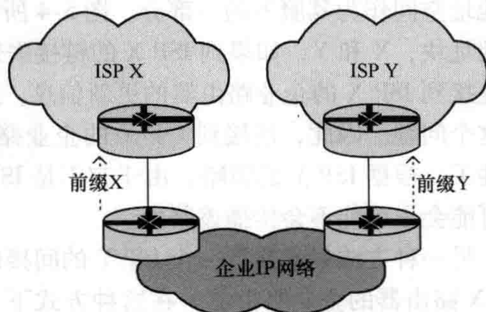



图 5-4　多重连接架构

- 链接冗余，提供了当连接中断事件发生时有继续可用的因特网连接。
- ISP 冗余，如果采用多个 ISP，可以减少由于某个 ISP 中断产生的影响。
- 在多个连接上分流因特网流量。
- 根据拥塞情况或路由不同应用程序的流量到不同的链接或不同 ISP，进行不同的路由，从而在策略与性能上获得收益。

多重连接提供了多个很有吸引力的好处，尽管它需要小心配置连接到每个 ISP 的路由器接口。就像图 5-4 所示，直接连接到它们各自 ISP 边界路由器的企业边界路由器，加入了一个外部路由协议（如 BGP），以通告其各自地址块的可达性（通过地址前缀）。因此，连接到 ISP X 的企业路由器将会通告由 ISP X 提供给该企业的地址空间的可达性。同样的，连接到 ISP Y 的企业路由器将会通告由 ISP Y 提供的地址空间的可达性。

这两个企业路由器也通过企业 IP 网络使用内部路由协议来进行彼此通信。在这种方式下，可能发现失去 J 与一个 ISP 的连接，这是令人感兴趣的地方。下面略过路由细节，来简要地说明这点。下面总结了最常见的多重连接的部署选项，运行中断的影响以及对 IP 地址空间的影响：

• 情况 1。两个或多个不同的物理链路连接到同一个 ISP。这种“多个附属”的架构提供了链接冗余，但并没有 ISP 冗余。如图 5-4 所示，两个 ISP 云将折叠成一个单一的云，但仍有从企业来的两个（或更多）链接。对于单个 ISP，前缀 X = 前缀 Y，所以这个从 ISP 分配的公共地址空间就会在所有连接上统一发布。

• 情况 2。使用供应商独立（Provider Independent，PI）的地址空间，通过两条或多条链路连接到一个或多个 ISP。PI 空间是独立于 ISP 而直接分配给某个特定组织的。如图 5-4 所示，两个连接的通告前缀再次相同，虽然可以将其记为前缀 Z 作为独立的 ISP 地址空间。如同情况 1，PI 空间可以被通告给所有 ISP，且需要在整个组织中进行分配。

• 情况 3。使用每个 ISP 的供应商聚合（Provider Aggregate，PA）的地址空间，通过两条或多条链路连接到两个或多个 ISP。在这种情况下，每个 ISP 将分配地址空间作为其服务的一部分。图 5-4 所示正反映了这一情况。有了两个独立的地址块，X 和 Y，如果到 ISP X 的链接失效了，由于内部路由协议的特性，根据连接到 ISP X 的企业路由器的更新信息，连接到 ISP Y 的企业路由器就会检测到这个问题。因此，连接到 ISP Y 的企业路由器现在就可以通告到前缀 X 的可达性了。根据 ISP Y 的策略，由于它不是 ISP Y 的地址空间而是 ISP X 的，因此它可能会也可能不会传播该路由。

另一种方法就是执行一个 ISP Y 的间接的 BGP 更新，而 ISP Y 其实是连接到 ISP X 路由器的企业路由器。在这种方式下，ISP X 路由器可能会被告知，通过 ISP Y 有一个替代路由能够到达企业地址空间。图 5-5 所示的多重链路中断恢复，说明了这两种可选的方法，前者展示了前缀 X 被通告给 ISP Y 的路由器，后者则展示了前缀 X 被通告给 ISP X 路由器[86]。

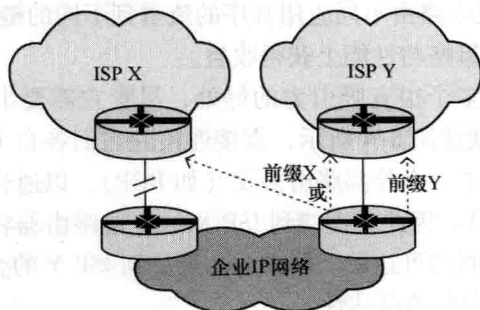



图 5-5 多重链路中断恢复

还有一个方法是部署 shim6[87]，它是协议栈中 IPv6 层正上方的一个“垫片”，用来用最佳下一跳来映射目标地址。启动了 IPv6 流量的终端设备在它们的

协议栈中将需要一个 shim6，以启用这种最佳的路由，这也可以实现重新路由以绕过某些断开的 ISP 连接。

每个 ISP 连接的 NAT 网关都启用了地址池，在 ISP 连接邻近的基础上，如从前缀 X 或 Y，将一个给定数据报的内部私有地址转化为一个公共地址。除非使用 NAT 网关，不然企业应该要实现边界路由策略，以最小化或阻止在内部地址主机间的流量被路由为通过 ISP 的流量。

## 5.6 IP 地址规划总结

本章讨论了制订你的 IPv6 地址计划的机制和技术。图 5-6 所示的 IPv6 地址规划总结了基本过程。作为你选择的部署范围中初步的整体发现与评估的一部分，其中一项成果应该是一份本书第 4 章中讨论的 IPv4 地址库存清单。这份库存清单应该根据你的网络拓扑来进行分层建模，以反映出拓扑级别、相关的分配、直到子网和独立 IP 地址的分配级别。该库存清单还必须列出 DHCP 地址池，以及为故障切换的准备的备份池、分割范围或共享子网等。

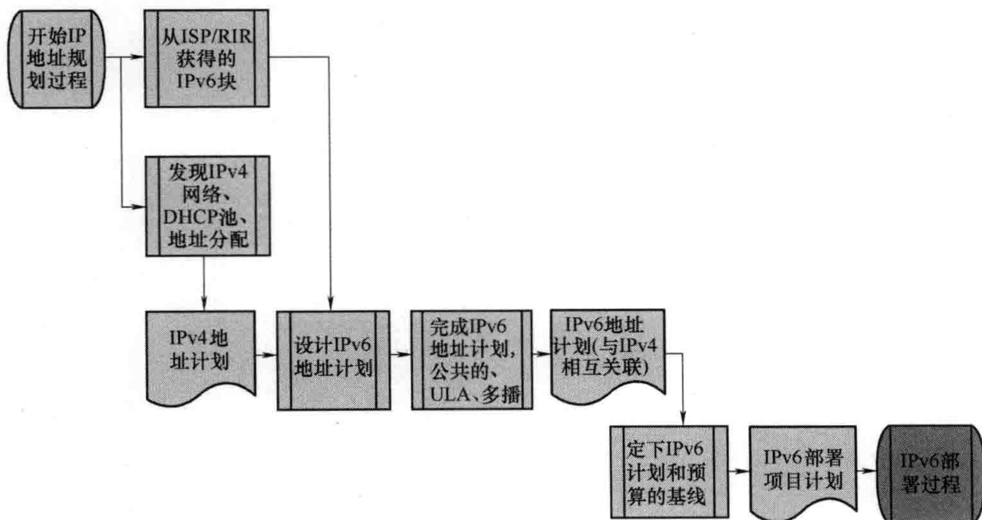



图 5-6 IPv6 地址规划

这项评估工作的最终结果，即 IPv4 地址计划，正如本章所描述的，为 IPv6 地址规划提供了基础。假定你计划在很大程度上保留当前的基本数据流和路由结构，那么这样将大大降低因为部署 IPv6 网络而破坏网络正常数据交换的可能。结合 IPv4 地址计划，你的 ISP 分配的 IPv6 前缀也可以相应地被应用和分配。在你的网络拓扑应用程序、管理和策略需求及管理授权需求的基础上，制定一个

IPv6 的分配层次结构并在对应的层次上进行稀疏或最佳或随机的地址分配。

对照 IPv4 计划来制订 IPv6 地址计划，对于保证完成分配的覆盖和确保分配层次符合网络的分层来说是有帮助的。计划还应该包括对 ULA 空间及多播地址需求（如适用的话）的考虑。一旦 IPv6 计划已经确定，该计划就会在部署实施过程中被参照遵循，以指导和保证路由器、每台设备、IPv6 子网/地址正常划分。DHCPv6 的准备也必须在主机地址分配和可能的基础设施投资时被考虑进去。其他与 IP 寻址相关的基础设施投资的可能是极小的，除非你需要一个 IPAM 系统来同时管理你的 IPv4 和 IPv6 空间。

# 第 6 章　IPv6 安全计划

如果计划在你的网络中引进一个新的网络层协议，必然会引起安全团队的重视。事实上，在你的网络中可能早已拥有了一些试图连接网络资源的 IPv6 寻址设备。因此，作为引进 IPv6 计划的重要组成部分，更新你的安全策略就非常关键了。本章主要从安全的视角来探讨 IPv4 和 IPv6 之间的差异，并突出了在更新安全策略时需要考虑的一些关键点。

## 6.1　好消息：IP 依然是 IP

如同 IPv4，IPv6 是属于 OSI 七层协议族中的一种网络层协议[88]。在一个 IPv4 的网络中使用 IPv6 并不会对网络层的上层或下层有潜在的安全影响。IPv6 本身并不比 IPv4 更安全或更不安全，但它与 IPv4 不同，必须从安全的角度加以考虑。因此，没有新的应用层、传输层、链路层或物理层漏洞被引入，也没有被消除。一般来说，以下的攻击类型应该继续包含在您的安全策略中：

- 物理安全性和访问。
- 未经授权的网络访问许可：通过第二层（如可扩展身份验证协议（Extensible Authentication Protocol，EAP），Radius/Diameter 协议）或者第三层（DHCP 或者 Spoofing）。
- 应用层，传输层，链路层或物理层攻击。
- 中间人攻击。
- 操作系统漏洞和攻击。
- 流量嗅探。
- 拒绝服务攻击（Denial of Service，DOS）和分布式拒绝服务攻击（Distributed DOS，DDOS）。

然而，相比单个协议栈，运行 IPv4 和 IPv6 双协议栈的设备更容易受到这两种网络寻址的安全漏洞的影响。对比 IPv4，业界对 IPv6 安全缺乏经验上的认识，使得上层攻击者直接攻击 IPv6 寻址比直接攻击 IPv4 寻址更加有效。应用本章讨论的 IPv6 安全策略将会帮助降低这些风险。

引入 IPv6 不是为了改变网络流量，但可能会增加 IPv6 单协议主机的流入或流出流量。用户设备最终产生的 IP 流量由用户的行为决定。如果用户引入一个 IPv6 设备，新的 IPv6 流量必然会产生，与用户使用传统设备产生的 IPv4 流量代

价是一样的。因此，除非在计划部署 IPv6 的同时恰巧重新设计主要网络（并不推荐这种做法），那么整体的 IP 流量模式在大多数情况下将会保持相同。

## 6.2 坏消息：IPv6 不是 IPv4

尽管好的一面是，引入 IPv6 带来的漏洞会大部分受限于网络层；但坏消息是 IPv6 不同于 IPv4，它具有必须被监督的独特特性，以及如果某些设置未被关闭，则必须详细审查。首先，IPv6 不是一项新技术，大规模的 IPv6 部署却还没有实现。鉴于迄今为止 IPv6 的部署仍处于较低水准，“臭名昭著”的攻击比高能见度、以 IPv4 为目标的攻击少。识别和处理这些攻击的实际经验有限。相反，伴随着一次次 IPv4 互联网攻击，应对攻击向量的经验不断累积，就能制定相应的 IPv4 应对方案和流程。

尽管 IPv6 有某些不同的东西，也引进了必须考虑的独特特性，但是基于 IPv4 的安全策略也能够适用于 IPv6。例如，考虑以下几点：

- 当 IPv4 使用 ARP 来将 IP 地址关联到链路层地址时，IPv6 使用 NDP——地址自动配置和重复地址检测需要这个协议。因此，ARP 攻击的缓解策略也应该适用于 NDP。
- IPv4 支持广播，而 IPv6 使用知名的多播来代替，如在 DHCP 中。
- 当需要逐跳的路由器来处理数据报时，路由器资源会被消耗。例如，在逐跳扩展报头内使用 IPv6 路由器警报选项。
- IPv4 的分段是在路由器上进行的，而 IPv6 是在主机上进行分段。
- 移动 IPv6 和移动 IPv4 相似，但也有很大的区别，本章将稍后讨论。
- 由于 IPv6 庞大的子网，使用暴力 ping 来识别主机相对来说是比较困难的。当然，如果您从 1 开始给主机分配地址并逐渐累加，将使得整个发现过程更加简单。
- 一般来说，在 IPv4 中，ICMP 能够被关闭；而在 IPv6 中，ICMP 是一个必需的协议，不能完全被关闭。
- IPv6 扩展报头的存在使得基本 IPv6 报头很短小，但会导致与报头相关的攻击。
- IPv6 协议栈软件的不成熟也有可能存在容易被攻击的漏洞。
- 正如本书第 3 章提及的 IPv4/IPv6 共存技术中，经常会包含多个交互的部件和复杂的操作，这也将带来一些容易暴露和被攻击的安全漏洞。

在接下来的 IPv6 安全策略部分将详细讨论这些问题。你应该将更新的网络层安全策略纳入 IPv6 策略内。正如你现在的安全策略，IPv6 缓解策略应该明确地被管理层记载、公示、认可，以及被网络用户所理解。接下来的几节将会着

重说明那些在寻求缓解潜在的安全漏洞下所建议的安全策略。

## 6.3　更新你的安全策略

首先考虑 IPv6 部署的安全隐患，其中有许多安全隐患已在本章指出；然后更新当前网络的安全策略文件，这点非常重要。一旦更新和通过了修订的安全策略文件，就要考虑防火墙、路由器、服务器、其他基础设施和终端用户设备对策略支持、识别存在的不足。这些不足可以通过升级、更换或记录有效的缓解策略来解决。这些升级计划和策略的实现都应该与整体的 IPv6 实施计划成为一体，与整体范围和部署阶段相匹配。本章的其余部分将讨论潜在的攻击类型和按基本网络安全类别分类的缓解方法。

## 6.4　网络边界的监控和入侵防护

抵御来自因特网攻击的第一道防线就是你的网络边界。传统面向因特网的架构，以路由器及包含可达外部资源和防火墙的“隔离区”（Demilitarized Zone，DMZ）为特点，这同样也适用于 IPv6。为了让 IPv6 能访问到你面向因特网的资源，就需要 IPv6 流量至少能进入 DMZ。内部用户与 IPv6 因特网目的地通信的必要条件就是允许 IPv6 流量进入内网。

根据以前记录的若干实例，来自因特网的攻击包括企图渗透、拒绝服务、劫持、阻碍性能的行为，以及通常破坏团体间 IPv4 网络基础设施的通信和服务。相似的攻击也可能出现在 IPv6 资源上。考虑到运营 IPv6 相对缺乏经验，众所周知的 IPv4 类型的攻击也可能会发生在 IPv6 节点上。实际上，就像之前提及的，攻击者可能会将 IPv6 视为既定的攻击对象的一个“后门”。所以，必须要采取行动来监视攻击，减少漏洞。与此同时，使用具有记录丢报能力的过滤器也是个好主意：审查丢弃的数据报不仅有助于找出潜在的攻击，还能识别出一些本应该被允许通过的“好”报，从而可能需要更新过滤配置。

### 6.4.1　IPv6 地址过滤

类似 IPv4 地址过滤，IPv6 地址过滤应考虑以下策略：

- 为了防止使用非法地址带来的欺骗，应丢弃接收到的未分配的 IPv6 地址空间的数据报。根据期望的过滤度，你可以在大范围分配水平简单地进行过滤，可以参考 http：//www. iana. org/assignments/ipv6-address-space/ipv6-address-space. xml[89]，或是明确 IANA 分配的每一个 IPv6 块 http：//www. iana. org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-addressassignments. xml[90]

• 拒绝具有以下定义的源地址和目的地址的数据报。

大部分防火墙都允许建立在“第一次匹配”基础上的策略规范。因此，假设想要允许全部的 IANA 分配的地址空间，那么你只需将此“2000::/3 allow”或“allow 64:ff9b::/96”语句添加到拒绝访问列表中；如果你计划支持本书第 3 章讨论的 IP/ICMP IPv4/IPv6 转换前缀，那么使用“deny all”；如果你更倾向于 RIR-level 过滤度，那么可以明确允许过滤度更精确的全球单播 2000::/3 访问。

考虑以下额外的安全策略：

• 丢弃入站目的地址或出站源地址为分配和公告的公共 IPv6 地址空间外的数据报。

• 配置单播逆向路径转发（unicast Reverse Path Forwarding，uRPF）过滤——通过配置路由来丢弃源地址不可达或是无法通过最佳路由的接口到达的数据报。这将有助于检测出可能有恶意的设备利用源地址欺骗发送的数据报。

• 尽可能采用深度报检测技术（Deep-Packet Inspection，DPI）来分析所有的 IPv6 扩展报头的一致性。这将在 6.5 节讨论。

• 考虑您的 IPv4 协议支持中“41”表示的是一个 IPv6 隧道数据报。详情参考接下来的 IPv4/IPv6 共存技术章节，讨论支持各类过度技术对安全的影响。

• 丢弃入站的目的端口号对应的服务或应用程序不支持外部接口的数据报，如 DHCPv6 对应端口 546 和 547。

• 使用具有身份认证功能的安全路由协议。

• 在外部 DNS 上只发布主机通过因特网可达的 DNS 记录，如 Web 服务器和邮件服务器，并考虑签署 DNSSEC。

• 通过定期检查日志，制定一个可用来发现异常的基准。当然，随着越来越多的用户加入 IPv6 网络，各流量模式可能需要花一段时间才能趋同。

### 6.4.2 ICMPv6 消息

ICMPv6 是 IPv6 里的核心部分。然而，它也有可能被网络攻击利用。因此，应该谨慎地考虑过滤处于边界的 ICMPv6 数据报（IPv6 下一报头值为 58），并拒绝错误消息类型进入，但除了那些包含你的确需要的消息类型的 ICMPv6 数据报。RFC 4890（即本书参考文献［91］）定义了 ICMPv6 过滤建议。在错误消息中，以下消息必须允许通过：

• 类型 1。目的地不可达。

• 类型 2。数据报太大。

• 类型 3。超时。

• 类型 4。参数问题代码 1 和代码 2（代码 1 为遇到未能识别的下一报头，代码 2 为遇到未能识别的 IPv6 扩展项）。

除此之外，RFC 不建议丢掉支持 ping6 应用的类型 128（echo 请求）和类型 129（echo 回复）。很多安全策略有意禁止入站的 IPv4 echo 请求，旨在阻止在一个给定的网络里检测有效的 IP 地址。在一个/64 子网内的 $2^{64}$ 个 IPv6 地址中，假设不是以::1、::2、::3…这样开始手工分配 IPv6 地址，那么检测主机将是非常困难的一件事。因此，找到给定 IP 地址的主机很可能就是攻击者确定攻击目标，或利用潜在“代理人”来发动攻击的第一步。

正常情况下，以下 ICMPv6 错误消息不应该被丢弃：

- 类型 3。超时代码 1（分段重组超时）。
- 类型 4。参数问题代码 0（遭遇错误的报头字段）。

关于移动 IPv6，也已经定义了确定的 ICMv6 错误消息。如果你的网络明确不支持移动 IPv6，以下 ICMPv6 消息应被过滤；否则以下类型不该丢弃：

- 类型 144、145。本地代理发现请求/应答。
- 类型 146、147。移动前缀请求/公告。
- 类型 154。移动 IPv6 代理路由器请求/公告。

RFC 4890 建议丢弃以下的链路本地流量，且无需特别注意，尽管你可能想在安全策略中明确定义过滤该类型，或是在防火墙验证该操作：

- 类型 133、134。路由请求/公告。
- 类型 135、136。邻居请求/公告。
- 类型 137。重定向。
- 类型 141、142。逆向邻居请求/公告。
- 类型 130~132、143。多播监听、发现监听查询，报告，完成，报告（v2）。
- 类型 148、149。SEND 认证路径请求/公告。
- 类型 151~153。多播路由器公告（Router Advertisement，RA），请求，终止。

至于那些只有在特殊情况下才使用的 ICMPv6 也应该丢弃，请仔细考虑以下几种类型：

- 类型 139、140。节点信息查询，响应。
- 类型 138。路由器重编号。
- 类型 100、101、200、201。私人实验。
- 类型 127、255。用于扩展的保留类型。
- 类型 150。用于实验。
- 类型 5~9、102~126、156~198。未定义（未分配）消息类型。

除非想运行一个低功耗有损的网络（Low-power and Lossy Network，LLN），否则也可以丢弃 ICMPv6 的类型 155。在 LLN 中，该类型用于支持 IPv6 路由协议的消息。

## 6.5 扩展报头

通常来说，路由器只需要分析IPv6基本报头、可能存在的逐跳选项报头、路由报头前可能存在的目的报头、路由报头，以及如果路由器是支持shim6的边界路由器时的shim6报头。然而，防火墙不仅涉及路由器需要处理的报头，还要解析上层协议报头来确定是否应该丢弃或允许通过该数据报。在一个给定的数据报中，除了可能会出现在路由报头前的上层协议报头里的目的选项报头，每个扩展报头都只应该出现一次。这应该是强制性的，以降低由于链接扩展报头带来的DOS型攻击的风险。有效的报头类型也应该被验证。

虽然扩展报头是有效且单一的，但攻击者仍有可能利用它们看似无辜的外表来发起攻击。例如，可以利用逐跳报头或目的报头的填充选项来创造一条隐蔽的控制渠道。填充选项原本是用来填补报头选项的整数字节边界，不该包含任何“数据”有效负载。攻击者可以利用路由器逐跳警报选项来达到拒绝服务或降低网络路由器的性能，因为此时路径上的每个路由器必须更加深入地检查数据报，这是一种很消耗路由器资源的行为。

除非你正在使用移动IPv6，否则包含路由报头的数据报就应该被过滤掉。类型0的路由报头已经过时，类型1主要用于实验，而类型2则支持移动路由选项。如有需要，至少且仅要支持类型2。

当一个数据报的大小超过了MTU，则需要使用分片报头。在IPv6中，只有终端才能对数据报分片。过滤分片数据报的一个问题就在于上层报头信息可能不会包含在第一个分片内，因此就需要对多个分片进行“有状态的”分析才能决定该数据报的命运。攻击者可能通过发送许多小数据报分片，来阻止以上处理过程，或者通过发送恶意信息或故意拒绝服务来阻止完全过滤。小于1280字节的数据报分片都值得怀疑，因为1280字节是IPv6 MTU的最小值。

除非是你的网络里（路由或主机）明确需要的报头，未知的扩展报头类型都应该被丢弃。在已存在的各类报头类型中，较新的逐跳选项和目的选项被定义为额外选项[92]。除了之前讨论的ICMPv6，可以借鉴当前IPv4的过滤实践。请记住，在IPv4网络存在的（分布式）拒绝服务攻击、缓冲区溢出攻击、跨站点脚本注入、SQL注入和其他上层攻击，也可能发生在IPv6网络中。

## 6.6 内部网络保护

许多组织通过建立一个很强大的边界和入侵检测系统来保护它们的内部网络。但是，内部网络保护也是必需的，以防止有意或无意的内部攻击，而且也

应该是避免外部攻击连入内网的第二道防线。

### 6.6.1 网络侦查

如果一个攻击者正在寻找一台主机来安装蠕虫或是后门代理，那么他的第一步通常是侦查寻找潜在的受害者。在 IPv4 中，如果允许 ping 扫描的话，则能提供一种简单且快速的方式来检测一个子网内活跃着的 IP 地址。在一台给定的主机上增加 TCP 端口放开（SYN）也能被用来尝试识别该台主机的操作系统。在 IPv6 中，根据 RFC 5157《IPv6 网络扫描的影响》（即本书参考文献［93］），通过 ping 扫描一个/64 子网需要花费 50 亿年。然而，RFC 5157 也指出，子网内这种难以预测的、稀疏的 IPv6 地址分配不是一个“安全方案”，尽管它能帮助降低蠕虫依赖网络扫描传播的有效性。另一方面，你从::1、::2 等开始手动分配子网地址的话，那么扫描的工作将会容易很多。

如果你正在使用没有私有扩展的 SLAAC，假设攻击者能够确定在网络上普遍使用的网卡制造商，那么他/她能够从 $2^{64}$ 个 IP 地址中极大地减少候选地址。一位前雇员或同事或许知道网卡的制造商，可以计算出 8 字节接口 ID 中的前 5 个字节，即能利用每个网卡制作商以太网 OUI 将扫描范围减少到 $2^{24}$ 内。6 字节 MAC 地址中前 3 个字节包含了 IEEE 分配的 OUI。倒数第 7 个最重要的位是修改 EUI-64 过程的一部分。然后追加 0xfffe，IDD 的前 40 位就可推导出来了。接下来只需找到剩下的 24 位。参考 RFC 5157 里“每秒扫描一个地址”的速度来扫描，则需要花费 195 天（或 6.5 个月）——依旧冗长的时间但却远远短于 50 亿年！

你可以使用 SLAAC 私有扩展[29]来避免这种确定的接口 DD 计算，以避免这种攻击向量。SLAAC 私有扩展阻止了猜测地址的行为，但却使得网络管理更具挑战性——在 IPv6 中，花精力将特定 IP 地址与特定主机关联起来的日子已经一去不复返了，但是使用私有扩展会为此增加难度。网络管理工具在取证和故障排除中可以通过使用定期轮询交换机和路由器来追踪 IPv6-MAC 地址的关联。

### 6.6.2 网络访问

除了发现已分配 IPv6 地址的已有主机外，另外一种本地网络攻击涉及攻击者在给定子网中获得自己的 IPv6 地址。这种形式的攻击可能始于在用户网络中远程安装僵尸程序的恶意尝试，例如，访客设备在会议室无意地接入网络，或其他各种“意外”的网络接入形式。2000 年中期，网络准入控制（Network Admission Control，NAC）的准则层是个很热门的话题。当时，几家知名的厂商提供了一系列有关于检测 IP 网络接入，限制未知或未认证的设备接入网络，以及加强设备扫描和病毒防护更新等的解决方案。

一些 IPv4 NAC 的设计原则也可以应用于 IPv6 NAC，但是 NAC 的实现却没

那么简单，通常需要协调两个或更多的网络设备。例如，第二层的NAC方案将涉及通过捕捉来自交换机的简单网络管理协议（Simple Network Management Protocol，SNMP）陷阱（trap）来检测相应交换机端口是否被激活（“link up”），然后由Radius服务器通过EAP协议向端口发起认证挑战。第三层NAC方案涉及一个DHCP服务器、一个DNS、一个认证服务器和一个设备扫描服务器。这些方案将在第8章详细讨论[66]。

当然，IPv6增加了另外一种IPv4不支持的地址分配形式——SLAAC。设备将自动配置IPv6地址，并进行地址重复检测，然后连入网络。如果该设备有意或无意地冒充路由器，它可能在链路上放出错误的路由公告，以致拒绝服务或拦截数据报。在你的交换机上利用RA-guard技术来阻断来自不是与路由器连接的端口的RA消息，能够减少这类攻击。RFC 6105详述了该技术。

网络中的设备在收到“合法”的设置“M”位的路由公告时，这样的设备支持使用DHCPv6来分配地址，但是大部分攻击者并不遵循任何规则。在处理一个特定源地址发来的数据报之前，DHCP的LeaveQuery功能可以用来验证该地址是否是由DHCP所分配的地址。该功能的运行方式是路由器一旦收到相连子网上设备发来的数据报，就发送一个LeaseQuery请求给DHCP服务器来确定：由MAC地址（通过DHCPv4）或IPv6地址（或DUID，如果知道的话）所标识的设备，是否有一个有效的租期。如果不是，那么该数据报将会被丢弃或是按照路由策略来处置。该方法肯定会增加路由器数据报处理功能的开销，但在某些环境中可能是值得的。另一种可选方法是在你的网络里利用SEND协议对使用邻节点发现协议的设备进行认证。

### 6.6.3 DHCPv6

DHCPv6具有和IPv4大致相同的漏洞，除此之外，攻击者接入你的网络后通过监听所有的DHCPv6（代理和）服务器多播地址——ff02::1:2（本地链路）或ff05::1:3（本地站点），可以更加容易地建立一个恶意的DHCPv6服务器，并处理请求（Solicit）或更新（Renew）消息。DHCP和DHCPv6都支持认证机制，但实际上由于初始配置的复杂性，IPv4上的认证机制很少被实施。如果DHCPv6服务器支持禁用这些多播地址，可以通过用DHCPv6服务器IPv6单播地址来配置中继代理，将DHCPv6服务器配置成“老式”的可达方式。通过周期性地轮询每个路由器上的邻接表来查看谁在网络上，并与之前的记录进行比较来检测网络上的新设备是否通过DHCPv6、SLAAC或甚至是手动配置了新地址。

### 6.6.4 DNS

如果一个网络没有配置为从可信任的DHCPv6服务器产生动态DNS

（Dynamic DNS，DDNS）更新，那么 DNS 更新安全则是另一个需要考虑的问题。当使用 DHCPv6 时，可以在 ACL 中定义允许更新 DNS 信息的 DHCPv6 服务器 IP 地址集合，除了如 IPAM 系统等管理系统外。仅允许固定的节点或地址集合对 DNS 更新，简化了 DNS 的 ACL 规范。当静态配置 IPv6 地址或使用 SLAAC 时，DDNS 需要更新的地址范围可能跨越整个网络，这将使得 DNS 易受到有害更新的攻击。首先要问的是，此类主机是否需要 DNS 条目。当然，对于 Web 服务器、打印机、应用服务器等用户使用 URL 或名字接入的主机来说，DNS 解析是需要的，DNS 中的条目是至关重要的。许多组织创建 DNS 条目至少是为了正向域名查询（主机域名到 IP 地址的查询）。反向查询（IP 地址到域名）是某些应用需要的，此时需要证明存在一个基本的安全措施。你可以选择占位符解析信息或是通配符子域（如 *.8.b.d.0.1.0.0.2.ip6.arpa）等变形 DNS 来最小化满足现有需求。使用 IPAM 系统（将在本书第 7 章讨论），可以为手动配置的 IPv6 地址自动化地创建和提供正向和反向的资源记录。

### 6.6.5　任播寻址

如果你正在使用任播寻址，如使用一个分配给多个服务器的公有 IPv6 地址，可以考虑配置服务器用这个任播地址作为返回数据报的源地址来响应一个到任播地址的入站数据报。这种方法适用于查询/响应的应用，但在面向连接的应用中可能更不可靠。正如 RFC 4942（即本书参考文献［95］）指出的，这种方法避免了服务器单播地址被检测到，特别是当请求被直接转发到一个如 DNS 这样的关键网络服务器时。

### 6.6.6　内部网络过滤

在其他内部网络漏洞方面，推荐将之前讨论过的以下过滤方法应用到内部路由器来保护网络边界。

- 丢弃接口上接收的源地址不在给定接口前缀范围内的数据报。
- 过滤没有在你的网络上明确定义使用的 ULA 地址空间。如果需要的话，拒绝表 6-1 给出的未分配或不合法地址空间产生的流量。
- 使用认证功能来保护路由协议安全。请注意，OSPFv3 提供了两种形式的认证：IPSec 只最初指定或结尾段认证（Authentication Trailer）。
- 丢弃具有路由报头的数据报，除非你支持移动 IPv6，而且只接收具有类型 2 的路由报头的数据报。
- 丢弃邻节点发现、重复地址检测及 SLAAC ICMPv6 产生的没有本地链路或未指明的地址（::/128），或不在 255 跳内的数据报。如果数据报从一个不同于本地网络的源地址发送过来，将会阻止这些消息的处理。考虑实现 SEND，如

果你的网络设备支持的话（不幸的是微软的操作系统暂时还不支持 SEND），否则，可以考虑创建更多较小的子网，而不是更少的较大子网，来减少暴露给本地网络攻击的范围。

- 除非在特殊情况下需要使用，这些 ICMPv6 类型的数据报应该被丢弃，仔细考虑以下类型。

√ 类型 139、140。节点信息查询，响应。

√ 类型 138。路由重编号。

√ 类型 100、101、200、201。私人实验。

√ 类型 127、255。用于扩展的保留类型。

√ 类型 150。用于实验。

√ 类型 5 ~99、102 ~126、156 ~198。未定义（未分配）消息类型。

- 定期检查日志以便制定一个基准用于发现异常。

**表 6-1 地址空间过滤建议**

| 过滤的地址 | 原　因 |
|---|---|
| :: | 拒绝不明确的源地址或目的地址 |
| ::1 | 拒绝回环的源地址或目的地址 |
| ::/96 | 拒绝 IPv4-兼容的源地址或目的地址 |
| ::ffff:0:0/96 | 拒绝 IPv4-映射的源地址或目的地址 |
| 2002::/24，2002:7f00::/24，2002:ff00::/24，2002:e000::/19，2001:6440::/26，2002:0a00::/24，2002:ac10::/28，2002:c0a8::/32，2002:a9fe::/32，2002:c000::/40，2002:c000:200::/40，2002:c612::/31，2002:c633:6400::/40，2002:cb00:7100::/40 | 拒绝非法 6to4 源地址或目的地址（分别对应的 IPv4 地址：0.0.0.0/8，127.0.0.0/8，255.0.0.0/8，224.0.0.0/3 (224.0.0.0/4&240.0.0.0/4)，100.64.0.0/10，10.0.0.0/8，172.16.0.0/12，192.168.0.0/16，169.254.0.0/16，192.0.0.0/24，192.0.2.0/24，198.18.0.0/15，198.51.100.0/24，203.0.113.0/24） |
| fe80::/10 | 拒绝外部或穿过内网的链路本地源地址或目的地址 |
| fec0::/10 | 拒绝不被接受的站点本地源地址或目的地址 |
| ff00::/8 | 拒绝多播地址作为源地址、拒绝带有非全网范围的目的多播地址却越界的数据报 |
| 2001:db8::/32 | 拒绝文件编制的源地址或目的地址 |
| 3ffe::/16 | 拒绝不被接收的 6bone 源地址或目的地址 |
| fc00::/7 | 拒绝越界的 IPv6“私有”（唯一本地）源地址或目的地址 |

## 6.7 网络设备的安全性考量

对于网络和终端用户设备来说，最后一道防线当然是设备自身的安全措施。如今，在 IPv4 中普遍使用的主机安全措施也同样适用于 IPv6，包括以下几项：

- 物理安全控制，如网络和应用的基础设施。
- 安全访问，通过本地控制台，SSH 或其他应用层网络协议。
- 密码管理策略。
- 终端用户设备安全策略和指令。
- 主机数据报过滤和防火墙。

√ 只允许支持服务（协议和端口）的数据报通过。

√ 拒绝具有表 6-1 给出的非法源地址的数据报。

√ 基于已定义类型及你的网络中使用的服务，来允许 ICMPv6 数据报入站/出站，正如在内部网络安全章节讨论的那样。

√ 拒绝具有错误或未知的 IPv6 报头的数据报入站/出站

√ 允许或拒绝在你的网络明确规定的隧道协议，如 6to4 和 ISATAP。

另外，如果正在使用的主机操作系统支持过滤 ICMPv6 数据报，检查关于全部或部分 IPv6 数据报的 ICMPv6 错误消息是否是真正需要由主机发送，如果不是则丢弃掉。ICMPv6 攻击可能伪造错误消息希望主机处理虚假的数据报来侵入设备。

## 6.8 移动 IPv6 安全

本书第 2 章介绍的范围包括了关于移动 IPv6 的基本概念。为了更好地理解该协议的安全隐患，那么先来了解通信路径是怎么在移动节点（Mobile Node，MN）、本地代理（Home Agent，HA）和通信节点（Correspondent Node，CN）间建立的。回想一下，MN 是移动 IPv6 设备，HA 是为 MN 与其本地地址（Home Address，HoA，如在 DNS 上公布的 MN 地址）提供连接服务的路由器，CN 是与 MN 相互通信的主机。

当 MN 在其归属网络时，CN 发送数据报给 HoA，HA 通过本地链路将它们传输给 MN。当 MN 漫游时，它将从它当前归属的网络上获取一个 IPv6 地址。这个地址就是转交地址（Care-of Address，CoA）。当 MN 漫游时将会通知 HA 它当前的 CoA 地址。使用这种方式，当 CN 发送数据报给 HoA 时，HA 就会拦截这些数据报并使用其 CoA 将它们隧道传送给 MN。返回的数据报则通过 HA 逆向路径发送回来。这种间接模式利用了 HA 到 MN 的所有的通信。还有一种更加有效且

直接的方式可以使 CN 和 MN 直接通信而不需要使用 HA 隧道传送。下面来了解一下这两种通信模式是怎么创建的。

### 6.8.1 移动扩展报头

移动扩展报头支持当前可达 MN 地址相关绑定信息的通信。移动报头（Mobile Header，MH）的类型字段值定义了消息的类型和相关参数：

- 绑定更新（移动报头类型值为 5）。通过 MN 发送给 HA 或 CN 来更新其 CoA 绑定。
- 绑定确认（移动报头类型值为 6）。通过 CN 或 HA 发送给 MN 来确认一个绑定更新或报告一个错误。
- 绑定错误（移动报头类型值为 7）。通过 CN 发送给 MN 指示一个移动有关的错误，其中包括目的选项报头的归属地址选项没有一个现有的绑定。
- 绑定更新请求（移动报头类型值为 0）。通过 CN 发送给 MN 请求在现有绑定上更新。正如接下来将要讨论的，移动 IPv6 支持返回路径可达过程（Return Routability Procedure，RRP），这能够使 CN 验证 MN 存在 HoA 和 CoA 地址，以便它可以可靠地接收来自 MN 的绑定更新的消息。

绑定更新是由 MN 发送至 HA 用来注册 MN 当前的 CoA 的。这些更新必须通过安全协议 IPSec 相互传达。如图 6-1 所示，其中消息类型由各自的 MH 定义；目的选项报头（Destination Options Header，DOH），路由报头（Routing Header，RH）和 ICMPv6 类型的值如图所示。其他报头参数，即源 IPv6 地址标记为 S，目的地址标记为 D，在每个消息里标注。

如图 6-1 所示，在移动节点获取到一个 CoA 后，它可能需要执行一个发现归属代理地址和前缀的任务。HA 地址发现允许一个操作者使用另外的一个 HA 地址，而不是将 HA 地址硬编码入每个 MN。请注意，在 HA 地址发现过程中没有强制的安全需求。MN 发送一个 HA 地址发现消息（ICMPv6 消息类型 144）给归属代理任播地址（{归属网络 IPv6 前缀}::7e）。HA 使用活跃的归属代理地址及 ICMPv6 消息类型 145 来回复。一旦 HA 被识别，MN 可以从归属代理请求一个前缀公告，该公告与固定网络的路由公告相似。请求前缀和公告的消息可以使用 IPSec 安全关联发送，可选择使用封装安全负载（Encapsulated Security Payload，ESP）。随后就是绑定更新过程，该过程允许 MN 向 HA 注册；绑定更新过程需要安全关联。

对于直接路由（换句话说，MN 和 CN 不通过 HA 路由）来说，其绑定更新和刷新过程被称为 RRP。这个过程（见图 6-2）需要 MN 通过发送相应的“初始化”消息，本地测试初始化（Home Test Init，HoTI）消息和转交测试初始化（Care-of Test Init，CoTI）消息，来初始化归属测试和转交测试。HoTI 消息使用

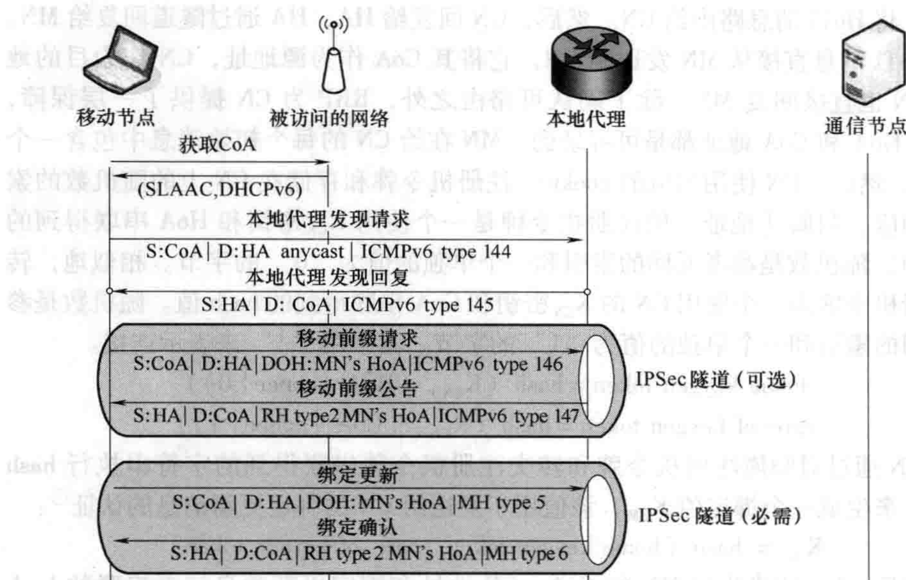



图 6-1　归属代理发现和移动节点注册

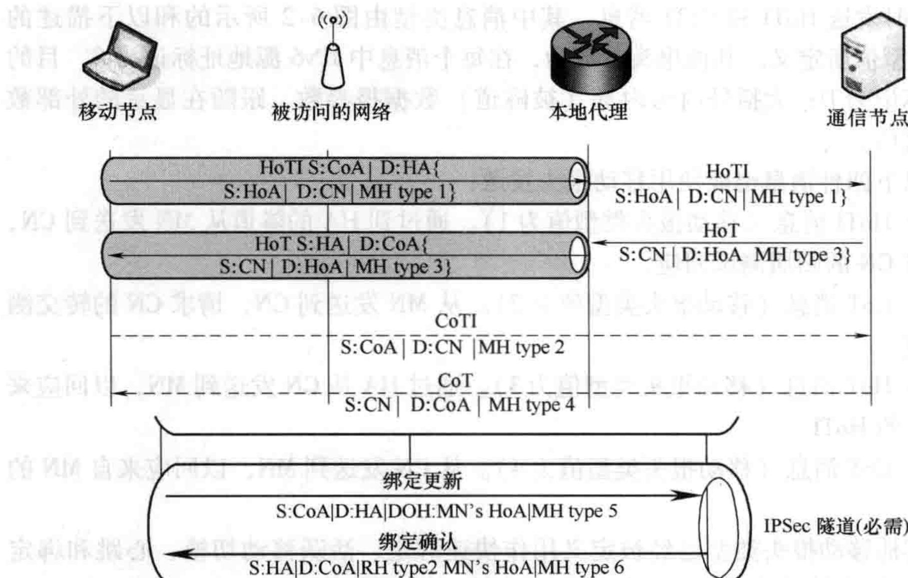



图 6-2　移动 IPv6 返回路由可达

MN 的 HoA 作为源地址，CN 的地址作为目的地址，并通过增加一个带有源地址为 CoA 和目的地址为 HA 的 IPv6 报头通过隧道方式传送给 HA。HA 除去隧道报

头后，将 HoTI 消息路由给 CN。然后，CN 回复给 HA，HA 通过隧道回复给 MN。

CoTI 消息直接从 MN 发送到 CN，它将其 CoA 作为源地址，CN 作为目的地址。CN 也直接回复 MN。除了确认可路由之外，RRP 为 CN 提供了一层保障，MN 的 HoA 和 CoA 地址都是可寻址的。MN 在给 CN 的每个初始消息中包含一个 cookie。然后，CN 使用相应的 cookie、注册机令牌和存储在 CN 上的随机数的索引来响应。归属 [地址] 的注册机令牌是一个使用 $K_{CN}$密钥和 HoA 串联得到的 hash 值，随机数是参考返回的索引和一个单独的值为“0”的字节。相似地，转交注册机令牌为一个使用 CN 的 $K_{CN}$密钥和 CoA 串联得到的 hash 值。随机数是参考返回的索引和一个单独的值为“1”的字节。其中用“|”来表示连接。

home keygen token = hash（$K_{CN}$，(HoA | nonce | 0)）

care-of keygen token = hash（$K_{CN}$，(CoA | nonce | 1)）

MN 通过对归属注册机令牌和转交注册机令牌串联得到的字符串执行 hash 算法，来生成一个绑定值 $K_{BM}$，该值用于发送给 CN 的绑定更新消息的认证[⊖]：

$K_{BM}$ = hash（home keygen token | care-of keygen token）

然后，$K_{BM}$用来生成 MN 的 CoA、CN 地址和绑定更新消息三者串联的 hash 值。HoA 与归属随机数索引和转交随机数索引标记一起在绑定更新消息内被传送。图 6-2 所示的移动 IPv6 返回路由可达性，说明了消息交换的过程。通常，MN 同时发送 HoTI 和 CoTI 消息，其中消息类型由图 6-2 所示的和以下描述的 MH 类型值所定义。其他报头参数为，在每个消息中 IPv6 源地址标记为 S，目的地址标记为 D；大括号内为内部（被隧道）数据报参数，跟随在显示的外部数据报值后。

以下四种消息也能使用移动报头发送：

- HoTI 消息（移动报头类型值为 1）。通过到 HA 的隧道从 MN 发送到 CN，来请求 CN 的归属测试回应。
- CoT 消息（移动报头类型值为 2）。从 MN 发送到 CN，请求 CN 的转交测试回应。
- HoT 消息（移动报头类型值为 3）。通过 HA 从 CN 发送到 MN，以回应来自 MN 的 HoTI。
- CoT 消息（移动报头类型值为 4）。从 CN 发送到 MN，以回应来自 MN 的 CoTI。

其他移动报头类型已经被定义用作快速绑定、活跃移动切换、心跳和绑定撤销。当前被定义的类型由 IANA 维护，详情如下：http：//www.iana.org/

[⊖] 为了删除先前的绑定，只需使用归属地注册机令牌来生成 $K_{BM}$进行身份验证，即 $K_{BM}$ = hash（密钥生成令牌）。

assignments/mobility-parameters/mobility-parameters.xml。

#### 6.8.1.1　路由报头类型 2

路由报头类型 2 支持直接路由从 CN 到 MN 的 CoA 的 IPv6 数据报。路由报头必须被路径上的每一个路由器分析。路由报头类型 2 可能只包含一个 IPv6 地址，一个 MN 的归属地址。因此，当数据报到达 MN 的指定 CoA 地址时，它将尝试处理报头来取出 HoA 和其固定地址。这意味着，MN 将会终止路由并处理数据报（即在 MN 的“本地接口”之间转发）。

#### 6.8.1.2　目的选项报头

移动 IPv6 利用标准的 IPv6 目的扩展报头来使漫游的 MN 将其 HoA 传递给接收者。该信息是由目的选项报头中的归属地址选项来传送的。

#### 6.8.1.3　移动 IPv6 消息流量小结

图 6-3 所示的移动 IPv6 消息流量，说明了返回路由可达过程后的绑定更新/确认过程和正常的 IPv6 通信。




图 6-3　移动 IPv6 消息流量

#### 6.8.1.4　移动 ICMPv6

正如本书第 2 章讨论的，移动 IPv6 规范已经定义了一些 ICMPv6 消息类型。为了方便，在此复述一下。

- 归属代理地址发现请求（ICMPv6 类型值为 144）。允许 MN 初始化动态 HA 发现。为移动的 HoA 前缀寻址归属代理任播地址，这使得移动时能够确定归

属网络中的 HA，如漫游时 HA 将被重新配置。

- 归属代理地址发现应答（ICMPv6 类型值为 145）。HA 回复一个归属代理地址发现请求，来确定移动 HA 的单播地址。
- 移动前缀请求（ICMPv6 类型值为 146）。使得 MN 能收集有关于其归属网络的前缀信息，如当归属网络发生重配置时。
- 移动前缀公告（ICMPv6 类型值为 147）。HA 通过这类消息公告当前归属网络前缀信息。
- 移动 IPv6 快速切换消息（ICMPv6 类型值为 154）。该类型既能被 MN 用来刺激路由器发送代理路由公告，也能为代理路由器为快速移动切换提供这种公告。

## 6.8.2 移动 IPv6 漏洞

移动 IPv6 已设计了如同非移动连接的安全防护。然而，仍然存在能被攻击者利用的漏洞。首先，如果你的网络不支持移动 IPv6 服务，那么应该明确禁止以下的相关 ICMPv6 消息：

- 过滤带有 IPv6 路由报头、IPv6 移动报头或是具有归属地址选项的目的选项报头（所有类型，包括类型 2）的数据报。
- 过滤类型值为 144 ~ 147 和 154 的 ICMPv6 数据报。

如果需要支持移动 IPv6，则要考虑以下漏洞：

- 通过模拟归属代理或通信节点来拦截和重定向通信的中间人攻击。
- 直接攻击移动设备操作系统、软件和信息。
- 防止拦截的通信安全。
- 拒绝服务攻击。
- 被访问网络（移动时）的安全策略。

### 6.8.2.1 恶意本地代理攻击

建立一个恶意 HA 使得攻击者能拦截发往本地网络的 MN 绑定信息，以及丢弃 HoTI 数据报来扰乱返回路由可达过程，从而迫使数据报通过 HA 路由。为建立恶意的 HA，攻击者必须要有一个本地网络的 IPv6 地址，这就需要先突破内部网络。

### 6.8.2.2 中间人攻击

恶意 HA 是中间人攻击的一种形式。一个或多个恶意 MN 可能会被利用来进行拒绝服务或至少通过多个假冒 MN 发起 DDOS 攻击来影响 HA 的性能。MN 向 HA 注册需要 IPSec 连接，这对于攻击者来说可能有点困难。假冒 CN 有相似场景，MN 可被引诱到一个虚假网站实施钓鱼来诱骗个人信息，尽管这与非移动 IPv6 连接没有什么不同。

#### 6.8.2.3　移动节点攻击

如同固定网络上的终端用户设备一样，MN 设备的操作系统、应用及储存在 MN 上的敏感数据极易被攻击。由于 MN“所有者”不能控制被访问网络的任何过滤策略，漏洞变得更加严重了。移动设备上应该提供某种形式的设备用作防火墙来降低设备受到攻击的风险。

#### 6.8.2.4　通信保密

大概绝大部分移动 IPv6 通信都至少涉及一种无线通信，因此遭窃听的概率大于固定网络通信。如果无线通信没有加密，MN-HA 通信的认证（可选择使用 ESP 功能）能够提供数据加密服务来减小漏洞。

#### 6.8.2.5　拒绝服务

如同其他设备，移动设备也可能成为 DOS 攻击的目标，或者它们会寻求安装僵尸软件来同时利用多个节点进行 DDOS 攻击。这些攻击的特点是向一个特定的 IP 地址发送大量的数据报，而移动节点有可能成为受害者。相对的，反射器或 smurf 攻击利用目标的 IP 地址作为虚假源地址，发送如 FTP 甚至 DNS 查询请求大量数据传输。速率受限的非移动设备或网络服务基础设施有助于转移此类攻击，虽然这一处理过程对预期的网络流量有性能的影响。

#### 6.8.2.6　被访问网络的安全性

当一个 MN 在漫游时，它必须从外部被访问网络获取一个 IPv6 地址。如果移动 IPv6 兼容的话，MN 向 HA 和 CN 注册这个转交地址。本地网络管理者管理 MN、HA 和本地网络，但是被访问网络通常是由其他实体（如 CN）来控制。攻击者可能企图假冒任何一个实体，来破坏或转移通信就像中间人攻击一样。坚持使用 IPSec 有助于保护移动 IPv6 控制消息。虚拟专用网络（Virtual Private Network，VPN）也可以用来保护被访问网络上的流量。

关闭移动 IPv6 控制消息类型的最低端的办法是你的网络不支持移动 IPv6。但是如果支持的话，使用适当的过滤器并使用 IPSec，是保护移动 IPv6 通信的关键技术。所以要检查你选择的 HA 和 MN 产品对 IPSec 支持的程度。

## 6.9　IPv4/IPv6 共存措施

正如本书第 3 章讨论的，有一些选项是用来在 IPv4 网络中部署 IPv6 的，尽管它们属于三大策略：双协议栈、隧道或转换。双协议栈至今还没有提出超出本书讨论范围的新要求；必须定义、文档化、实施和管理针对 IPv4 和 IPv6 通信安全策略。另外一个附加的管理要求是利用所有的地址——IPv4 和 IPv6，来追踪设备。对于可追溯、审计、网络访问追踪、取证分析和故障排除任务来说，这是必需的。

## 6.9.1 安全隧道实施

在 IPv4 中，隧道可当作攻击载体，将恶意的数据报封装在无害的隧道报头内。只分析外部报头的过滤器极易受到这类攻击的影响。由于除了 IP 地址格式匹配外，没有任何预先的隧道配置，只要基本数据报参数是符合格式的，那么隧道数据报通常都会被接收并处理，自动隧道技术也极易受到攻击。大部分配置或手动的隧道都能够通过检查协议字段是否是“41”来过滤此类 IPv4 数据报，协议字段为 41 则表示封装了一个完整的 IPv6 数据报。

### 6.9.1.1 6to4

6to4 在 IPv4 上为 IPv6 提供了以使用2000::/16 地址空间为特点的自动隧道。如果要拒绝 6to4 的流量，可以丢弃 IPv6 源地址或目的地址为 2002::/16 的数据报。6to4 路由器能够将 6to4 IPv4 数据报路由到 6to4 中继路由器，该路由器用作 6to4 IPv4 域与纯 IPv6 间的网关。RFC 3964（即本书参考文献［97］）详细说明了 6to4 的安全考量，并为 6to4 路由器推荐了以下过滤策略：

● 表 6-1 也给出了应该被过滤的私有或不合法的 IPv4 地址空间对应的 6to4 地址。

● 过滤带有外部源 IPv4 地址的数据报，该源 IPv4 地址并未被嵌入到 IPv6 源地址（即 6to4 的网络前缀中）的 IPv4 地址。

● 过滤目的 IPv6 地址是本地链路地址或是 IPv4-映射地址的数据报。

● 配置中继路由器公告只在来自/发往 IPv6 域的中继数据报中使用 2002::/16，而不用在 6to4 节点间的路由（6to4 中继路由器应该总是 6to4 隧道的终点）。当这样配置后，过滤从中继路由器收到的源地址不同于中继路由器的 6to4 地址（映射到外部 IPv4 源地址）的数据报。

● 过滤 6to4 目的地址不属于你公告的 6to4 前缀的数据报。

● 对于中继路由器，丢弃目的地址不是其中继路由本身的 6to4 数据报。

### 6.9.1.2 站内自动隧道寻址协议

ISATAP 规范明确指出了使用 IPv6 IPSec 是有必要的。它还推荐一般的 IPv4 入口过滤，特别是 IPv4 报头字段为“41”的数据报——表明封装了 IPv6 数据报，来减少在一个 ISATAP 链路中数据报注入的漏洞。ISATAP 路由器上的过滤配置也应该阻塞 ISATAP 路由器间的纯 IPv6 数据报来防止数据报被往返传递（ping-pong），并阻塞 IPv4 源地址不在本组织合法 ISATAP 用户范围内的数据报。

攻击者控制 ISATAP 客户端用来传输 IPv6 数据报的首选路由列表（Preferred Router List，PRL）能将 ISATAP 流量重定向到攻击者的路由器上。PRL 应该保持更新，而相关主机名为“isatap”的 DNS 条目也应该经常检查。

#### 6.9.1.3 Teredo

Terodo 被设计用来在你的防火墙上开个洞，使得 IPv6 数据报隧道能通过防火墙。RFC 4380，即 Teredo 的规范，指出了一些安全漏洞和建议缓解方案，见表 6-2。

**表 6-2 Teredo 漏洞和缓解方案**

| 漏 洞 | 缓 解 方 案 |
|---|---|
| 防火墙漏洞 | 限制一些本地链路服务<br>实施本地（主机）防火墙<br>使用 Teredo 隧道内支持的 IPSec |
| 作为中间人拦截和欺骗路由公告的 Teredo 服务器 | 在 Teredo 服务器和客户端间进行随机数认证。攻击者必须在“路径”上，因此除了拒绝客户端服务，攻击将很难进行 |
| Teredo 中继欺骗 | 使用 Teredo 直接 IPv6 连接测试流程，该流程由随机数认证保护 |
| 端到端的保密性 | 使用 IPSec 来阻止欺骗和窃听 |
| 拒绝 Teredo 服务 | 使用本地主机 Teredo 中继，要求身份认证、停用本地发现，以及部署不同的 Teredo 服务器，这样有助于减少各类拒绝服务类攻击 |
| 使用 Teredo 服务器拒绝目标主机服务的反射器攻击 | 通过观察已知流量的规律和语义，可以将反射流量识别出来 |

### 6.9.2 安全转换实施

转换网关处理所有需要转换的数据报，这可能是根据你的部署计划所产生的所有面向因特网的流量。DOS 攻击的防护和非期望 IP 源地址和目的地址的过滤，都应该实施。鉴于攻击者有能力修改 IPv6 前缀来破坏通信，用来转换的 IPv6 前缀的配置必须要足够安全。此外，RFC 6052 推荐支持对 IPv6 可转换地址中嵌入的 IPv4 地址的过滤[48]，就像对原生 IPv4 地址的过滤一样。

值得注意的是，尽管 ESP 隧道模式的数据报能够被转换，但 IPSec 身份认证报头不能使用 IPv4/IPv6 跨协议转换。只有 checksum-neutral 地址被使用时，ESP 传输模式才能成功在 IPv4/IPv6 跨协议转换中使用[47]。

## 6.10 小结

在基础设施、IP 地址和网络管理中，定义一种安全策略来支持 IPv6 实施是

一个重要的规划步骤。应该根据本章讨论的计划支持的IPv6功能，更新当前的IP安全策略。这些更新策略则需要与当前安全基础设施和系统一一对应。需要确定差距，然后通过更新、购买或替换软硬件，或是通过确认不足和记录暂时解决方法，来减少风险。

图6-4所示的IPv6部署的安全计划，说明了本章提出的发现和所需的IPv6安全策略比较的过程。任何缓解步骤必须被记录，并将财力和人力上的潜在预算编入，最后反馈到总体部署项目计划中。

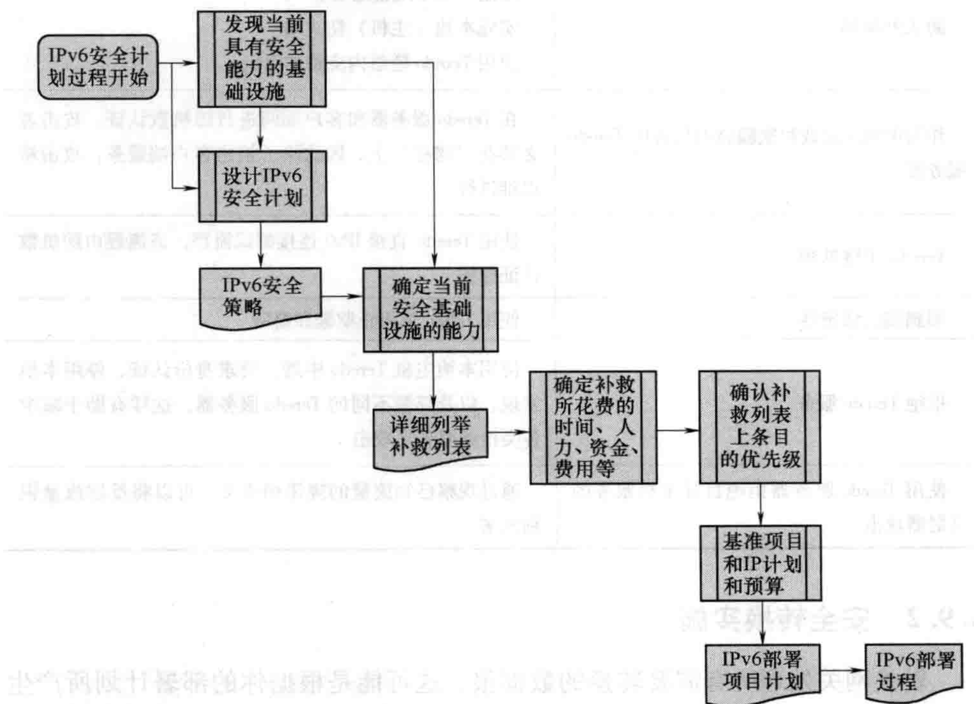



图6-4 IPv6部署的安全计划

# 第 7 章 IPv6 网络的管理计划

从技术上来说，本章应该叫做 IPv4/IPv6 的管理计划，因为你很有可能在保持 IPv4 网络的基础上部署 IPv6 网络。正如已经知道的，引入一个新的网络层协议 IPv6 到原有网络中需要考虑非常多东西，包括新协议的地址分配、新协议对于每一个设备的可达性、新协议的应用支持和新协议的安全问题。管理 IPv6 部分的网络是部署 IPv6 计划中一个与其他部分同等重要的组成部分。不管怎么样，如果 IPv6 的某个部分可以被部署，但是不能被合理地配置或监控，那么很少组织会实施这种部署。这也是必然的结果，因为网络管理是保证网络运作的一个非常重要的功能。除了少量本地设备组成的网络能在每一个接口上通过蛮力方式配置之外，网络管理通过给分布式网络提供一个集中视角产生了很大的价值。因此，通过对检测到的网络事件进行关联和安全优先排序，理顺了资源需求和提高了效率。

## 7.1 管理模型

最经常使用的网络管理框架是用于网络管理的 FCAPS㊀模型。IT 基础架构库（Information Technology Infrastructure Library，ITIL）是一套流行的管理企业 IT 基础设施的指南。由英国的政府商务办公室（Office of Government and Commerce，OGC）开发的 ITIL，是把 IT 组织看成企业的服务提供商的一个最佳实践框架。同时，它整合了一些和 FCAPS 相似的功能，虽然它的多数功能是面向服务的。

FCAPS 模型在网络管理的实践中包括如下几个主要功能：

- F，指故障管理（Fault Management），包括监视和检测网络故障，并能诊断、隔离和解决这些故障。具体来讲，需要对一些像路由器、服务器和交换机等网络单元进行监视并发现故障或运行中断，以及像 DHCPv6 和 DNS 这样的网络服务同样能够被监视。
- C，指配置管理（Configuration Management），包括对路由器和交换机这样的网络单元进行精确配置和备份，以及网络服务。其中至少包括网络层的应用和数据库服务器的配置和备份。对网络单元精确和及时的配置能减少错误产生，

㊀ FCAPS 是由 ITU 标准 M. 3400 定义的，是数据网络管理中的通信管理网络（Telecommunications Management Network，TMN）框架的一部分。

也能减少在转换管理窗口时的延时。

● A，指计费管理（Accounting Management），包括基于业务限定或客户授权范围内追踪和监控网络资源的使用情况。访问控制策略（access control policies），涉及业务参数的网络使用，以及对服务水平协议（Service Level Agreement，SLA）监控都在计费管理范围内。

● P，指性能管理（Performance Management），处理追踪网络单元和网络服务的性能和资源利用的问题。具体来说，追踪网络资源使用情况、网络/服务器的性能和数据流都在性能管理的范畴之内。

● S，指安全管理（Security Management），包括保护网络和网络用户的信息安全，提供访问控制、审计日志和安全违规检查。之前的章节已经对安全这一块进行了清晰的阐述，所以本章将不再讨论这个话题。

将会以 FCAPS 为背景来讨论 IPv6 网络管理，以适当的篇幅讨论包括像资产追踪（inventory tracking）、地址管理、配置信息数据库和发行与修改管理在内的功能。实际上，下面会从“配置”功能，特别是资产和地址管理开始讨论，因为这些功能对于定义 FCAPS 中其他功能的范围至关重要。举例来说，你的路由器资产是包含在进行故障或性能监控的一组设备中的。

## 7.2 网络管理的范围

在 IPv6 部署计划所涉及的全部范围内，你进行 IPv6 部署的网络部分由对 IPv4 及 IPv6 信息可以管理和追踪的部分组成。一定要从网络管理角度来考量在部署范围内的每一个基础设施或用户终端、应用程序和交换网络组件。这些都必须要从评估你的目标范围的当前组成部分开始。

### 7.2.1 网络库存

网络库存（network inventory）是指在你的网络范围内或目标范围内所有的路由器、交换机、存储点、通信服务器、打印机和用户终端设备的资产。这个库存是 IPv6 部署计划中的重要组成部分，可以用于确认设备、操作系统和应用究竟是本身就支持 IPv6 还是需要修改或升级以支持 IPv6（正如本书第 4 章所讨论的）。对每一个网络单元的设备，需要记录以下信息：供应商、制作模型（make model）、操作系统和版本、功能或应用、设备的硬件模块和网络接口。设备标识信息也需要被记录，这些信息包括：主机名、DHCP 唯一标识符（DHCP Unique Identifier，DUID）、每个接口的 MAC 地址、每个接口的接口关联标识符（Interface Association Identifier，IAID），每个接口的 IPv4 和 IPv6 地址。设备访问信息应该要安全地维护好，包括像 SSH 登录名和密码，以及 SNMP 团

体名和密码等信息。

### 7.2.2 IP 地址库存

IPAM 和库存管理功能在把 IPv4 和 IPv6 地址分配给网络中的合法设备时存在交织。此外 IPAM 也可以帮助发现网络中使用 IP 网络的设备，并且可以帮助识别网络中的新设备和移动过的设备。同时 IPAM 通常情况下比库存管理的范围来得广，因为它的功能还包括分别使用相应的主机域名和 IP 地址池来配置 DNS 和 DHCP 服务器。

实际上，在根据拓扑结构为分配地址而映射层次化地址块的时候，IPAM 功能和网络拓扑结构相吻合。正如在本书第 5 章地址规划中讨论到的 IPv6 需要分层地址分配。运用有约束的 IPAM，使得单个地址分配、路由通告和 DHCP 地址池可以遵守分层地址计划，在保持网络地址一致的同时简化了地址分配过程。总之，资产管理功能提供了一个含有相关属性的网络设备统一库存，这些属性包括与 IPAM 功能关联的 IP 地址，IPAM 功能是把 IP 地址分配给子网和动态地址分配的 DHCP 服务器，并附带相关的 DNS 域名信息。

通常情况下，设备资产在 IPv6 部署之前和之后应该基本相同。即，IPv6 的配置不需要新的硬件，虽然有的时候可能会需要对一些传统设备进行更换。另外，有的人也许会在初期通过增加部署额外的一些硬件设备来暂时隔离 IPv6 服务器，不过这个只是一个可选项。因此，一个在 IPv6 部署以前已经维护准确的设备和 IP 库存仓库（如作为评估阶段的结果）的组织，应该已经准备好进行 IPv6 的部署了，而且它们应该能够对每个 IP 寻址计划进行 IPv6 子网和地址分配来补充资产和 IPAM 仓库。

### 7.2.3 管理网络

因为一些或者全部路由被配置到 IPv6 网上并不意味着管理信息一定要通过 IPv6 来传输。例如，轮询 IPv6 相关的管理信息库（Management Information Base，MIB）可以用 IPv4 来完成，正如可以用 IPv6 来轮询 IPv4 相关的数据。如今在 IP 网络主要的管理应用协议，如 SSH、Telnet、TFTP、syslog、ping、traceroute 和 SNMP，都支持 IPv6。虽然轮询信息和检测警告可以使用任何协议，但是连通性和可达性的测试，像 ping、traceroute，或者网络测试工具，像 dig（DNSlookup 工具），应该要在对应的传输协议上运行才能模仿客户体验。

## 7.3 简单网络管理协议

SNMP 是 IP 设备实际上的网络管理协议，这些 IP 设备至少包括像路由器、

服务器、打印机等基础设施设备。SNMP 定义了对被管理设备的配置和信息检索的通信协议。它的基本结构包括一个网络管理者（network manager）。这个网络管理者使用基于 IP 的 SNMP，通过各个被管理设备上运行的代理来与一个或多个被管理设备通信。这个代理可以从被管理设备上访问相关的数据，并通过 SNMP 与网络管理者交互。被管理设备上这些可以进行配置和/或被收集的信息，是通过设备所支持的各 MIB 中的数字和类型所定义。

SNMP 为被管理信息（managed information）定义了一个整体的信息层次，而各 MIB 定义了信息对象的结构，或一组给定度量的变量。MIB-Ⅱ是目前正在使用的 MIB 版本。你可以认为 MIB 是一个数据库的表，而表的每一行是一个对象标识（Object Identifier，OID），每一个 OID 是由自己在标准 SNMP 对象分层中的位置唯一数字标识的。这让任何厂商管理系统能通过一个标准的 OID 值来访问已知变量。

SNMP 标准最初被开发出来的时间比 IPv6 作为一个网络协议标准被开发出来的时间要早。因此 MIB 定义的 IP/ICMP、TCP、UDP 只对 IPv4 的地址格式和信息有效。在 IPv6 标准被提出以后，新的 MIB 就对 IPv6、ICMPv6、TCP 和 UDP 进行了定义。此后，标准就演化成能够在传输过程中支持 IPv4 和 IPv6，并且，MIB 的数据结构也变得能够支持 IPv4 和 IPv6，而这一切都是从支持 IP 地址数据结构开始的。在旧的 MIB-Ⅱ版本中，一个 IP 地址文本习惯上是约定为 4 个八位字节串，这也是所有和 IP 地址有关的对象变量的标准习惯。RFC 4001（即本书参考文献[98]）定义了互联网地址类型（InetAddressType）和互联网地址（InetAddress）对，分别标识 IPv4 或 IPv6 的协议类型和对应的地址。这样，对象属性中表达的 IP 地址就可以通过类型和实际地址确定了。

除了统一的 IP 地址文本习惯，SNMP 现在对 IP、TCP、UDP 和 IP-FORWARD（路由信息）支持通用的各 MIB，而非单独的 IPv4 和 IPv6 版本。虽然有一些设备可能会含有特定供应商自己定义的和特定应用定义的 MIB 信息，但这整体上对节省配置 SNMP 管理器和“IP 地址”管理的时间十分有益。

### 7.3.1 配置管理

网络资产和 IP 地址计划决定了管理的范围。正如前面章节关于安全的叙述，周期性地扫描网络来发现新的或移走的 IP 地址或设备是一个好主意。对于每一个被管理设备，配置可以是通过使用很多不同的配置管理工具来管理的。这些工具必须支持你所选的传输协议——IPv4、IPv6，或者同时支持两者，而且必须能配置 IPv6 参数。

如果你计划使用 IPv6 作为传输协议，定义好 IPv6 地址分配机制是很有必要的！你可以使用 SLAAC 和 DHCPv6 中的一个或两个；也可以考虑使用隐私选项，

虽然这并不能用在那些需要使用可靠或固定地址的基础设施上。使用 IPAM 系统，不但可以减少在同时进行 IPv4 和 IPv6 地址空间管理时出现错误，也可以在大多数情况下自动为 DHCP 服务器和 DNS 提供相应的地址、域名和资源记录信息。不论如何，每一个网络资源仓库的入口都应该连接到它已分配的地址和分配方法，以及其他相关需要的配置和初始化参数上。

配置或者查看分配给网络设备上的 IPv6 地址要求管理系统支持 IPv6 地址的解析和显示功能。这和别的使用 IPv6 进行通信和显示的应用一样。所以，当评估你当前的网络管理工具能力的时候，这是一个非常重要的标准。

### 7.3.2　故障管理

故障管理必须监控网络连接和设备、告警，和排除故障解决问题，以使网络顺畅运行。SNMP MIB 轮询（polling）或者 SNMP 陷阱（trap）通常被用来监控设备状态，侦测告警状态。这些被监控和捕捉的各 MIB 和陷阱应该被文档化，而且应该为每个被监控设备所支持。正在使用的或计划使用的网络管理系统，也必须实现对这些 MIB 和陷阱的可视化和陷阱处理过程进行支持。

其他供应商定制的管理通告格式应该要支持 IPv4、IPv6，或者同时支持两者。一旦一个监控参数的变化超过了它的“正常”范围，必须采取主动行动来识别这种变化的原因，并采取任何正确而必要的行动来避免更大的网络问题。当然，所谓“正常”是通过对给定变量长时间的观察并计算出它的平均值和可以接受的方差来决定的。根据被监控的变量及其在整个网络中的影响，度量阈值可以是严格或渐变地提供信息（info）、警告（warning）和严重（critical）等不同等级的通告的。

在刚开始部署的时候，特别是从安全监控的角度来说，为了“超管理”早期的 IPv4/IPv6 运营，阈值可能会被严格设置在期望的平均值附近。如本书第 6 章所说，防火墙必须有详细的日志记录和分析，以识别所有破坏策略的行为。在安全和通信两方面，在基本顺畅运行数月之后，可以放宽阈值，但仍然需要临时的“抽查”来复检。

监控网络占用情况，或者监控有哪些拥有自己特定 IP 地址的设备在网络上，对验证地址计划和地址分配的合理性，以及发现潜在可疑设备或欺骗设备，是非常重要的。周期性地轮询（polling）路由器的 ARP 表或者 IPv6 邻节点表，为检测上述信息和状态提供了方便。

故障排除过程的文档应该明确，当超过阈值或 IPv6 相关或 IPv6 寻址设备相关的告警被检测到时的相关处理步骤。双协议栈设备在一个协议上不可达，可能在另外一个协议上是可达的，这为检测地址分配、路由或设备 IP 的协议栈等问题提供一个十分有效的方法。

### 7.3.3 计费管理

计费管理包括对网络资产的追踪。计费管理中很大的一部分任务是处理对设备、IP 地址分配、用户和应用的追踪。由于 IPv6 子网中的地址空间很大，把地址追踪技术转换到 IPv6 上是很有挑战性的。轮询路由器 SNMP 的“ipNetToPhysical”表为从一个中心地点追踪各个 IPv6 地址分配提供了一个好方案。像 nmap 这样的网络探测工具也可以通过实施多播 ping 来引发响应，或其他方式进行被动发现。不管收集机制是什么，把收集结果报告给中心 IPAM 系统可以帮助发现子网中新的移走的或者移动的设备。如果你的 IPAM 系统支持这些方式的发现功能，你已经可以追踪 IPv6 设备了。否则，应该向你的供应商询问最新的路线图。

### 7.3.4 性能管理

很多性能管理系统使用 SNMP 进行数据收集、数据聚合和数据报告。所以，考虑到之前讨论过的 SNMP 能够同时支持 IPv4 和 IPv6，应该能够很轻易地让性能管理系统支持 IPv6。也许实施这个系统最大的困难在于用户接口的数据表示，这个数据表示混合了 IPv4 和 IPv6 流量信息。网络探针这时候也可以用来“监听”IPv6 流量和相关的 TCP 流量控制和 ICMPv6 信息，以发现流量模式和潜在问题。

## 7.4 方法和过程

多数的组织会为预配置、监控、测试及对计算或网络设备诊断的标准化方法和过程（Method and Procedure，M&P），建立文档。随着 IPv6 的部署，这些文档需要更新。虽然部署 IPv6 就意味着引入另外一个可能在网络中出现错误的东西，但是它也为指定设备的异常处理提供了另外一个可选的网络层路由。设备在一个协议上可达而在另外一个协议上不可达，可能说明是路由或者是设备协议栈问题。在更新 M&P 文档时要考虑的关键点中，应注意以下几点：

- 分配地址块和子网的过程。特别是在为一个单一 IPv6 子网或一个 IPv4/IPv6 双栈环境，分配地址块和子网时。
- 如果使用的是 SLAAC 或 DHCPv6，准备和配置新设备的过程可能是不费力的。不过对于那些手动设定地址的设备，准备过程需要记录 IPv4 和 IPv6 地址的分配情况以及相关的参数。
- 异常检测指南应该包含 IPv6 相关统计信息和告警，以及用于隔离和解决异常的所需要采取的行动。

● 故障排除步骤应该要包括同时或分别通过 IPv4 和 IPv6 与一个指定设备通信的步骤。

● 计费功能中的发现模块应该要包括对 IPv6 邻节点表轮询，或者如 nmap 这样的发现工具的使用。分析不同地址分配方式时必须要考虑到 SLAAC 中的隐私扩展和 DHCPv6 的使用。而对未授权设备的检查功能需要通过对发现的主机信息与授权的 DHCP 唯一标识符和/或 MAC 地址的知识库比较来实现。

● 性能管理过程必须要包括 IPv6 MIB 统计数据和性能结果报告。

● 安全系统监控过程必须包括本书第 6 章所述的 IPv6 过滤指南，记录日志检查步骤来帮助发现可能的攻击，并验证安全策略设置。

设备部署的测试阶段为改善 M&P 文档提供了一个好机会。

## 7.5 小结

管理 IPv4/IPv6 网络的基本步骤和管理纯 IPv4 网络一样，但是必须考虑两种协议的所有网络管理方面，包括配置管理、故障管理、计费管理和性能管理。现在很多网络管理工具都同时支持 IPv4 和 IPv6 的信息收集，而不是信息的传输。确定自己网络管理设备是否有这种能力请参考本书第 4 章所描述的分析过程。根据所需的升级和已有资源，一个优先修正措施列表就可以添加到整个 IPv6 部署计划中去。IPv6 网络管理计划过程如图 7-1 所示。

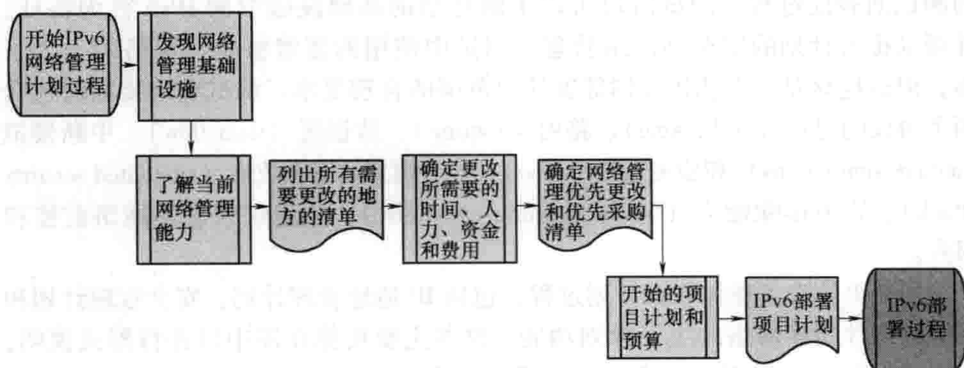



图 7-1　IPv6 网络管理计划过程

# 第8章 部署管理

## 8.1 整体计划

在之前章节叙述的计划工作的基础上，现在可以开始准备把所有计划整合起来了。这其中包括了地址计划、基础设施的修改或增添的计划，安全策略升级计划和网络管理计划。这个整合让你有能力找出所有跨职能的依赖关系，并且可以标出所有要做任务的时间表和资源可用性。在一些组织中，防火墙管理者（firewall person）和网络管理员（network manager）是相同的，那么与这两个角色相关的功能也就必须考虑到这个因素。但是，在大公司中，可能会由不同团队来负责不同功能，所以很可能很多工作都是并行执行的，因此需要各个团队间更多的合作来管理各种依赖性，并处理各种意外事故。

整合所有组成计划也能够明确部署测试的用例。即使没有明确要求，也建议最好在生产部署之前进行测试。测试计划应该根据部署 IPv6 带来的基础设施、地址计划、网络管理和安全策略的必要变化，记录与网络需求变动相关的测试的验证过程。测试阶段可以了解计划的基础设施中的 IPv6 行为特征，并帮助提升计划的安全和网络管理。测试中使用的要素将包括网络的一个子集，但这应该是一个能代表你部署计划范围的合理复本。测试应该要同时包含两个协议的寻址（addressing）、路由（routing）、数据流（data flow）、中断模拟（outage simulations）和检测（detection），包含模拟的安全攻击（simulated security attack）、故障排除能力（troubleshooting capability），还应包含总体网络监控和报告。

图 8-1 给出了整体联合计划过程，包括 IP 地址管理计划、安全管理计划和网络管理计划等网络的主要计划功能。这些主要功能在图中以并行形式说明，汇聚输出到“确定项目基线和 IP 计划与预算（baseline project and IP plan and budget）”过程中。整个计划的定购任务、识别依赖关系和购物清单条目在这里汇聚。一个整体的预算修改也会基于购物清单及不同资源的需求时间来准备。委任一个项目的经理来负责定期的项目状态会议、进度检查、启用意外事件计划和减少组织之间的问题。召集代表每一个功能范围的项目成员参加状态会议，提供自上次会议以来的活动情况、新的问题或花费，以及其他成员任何信息或资源需求的报告。

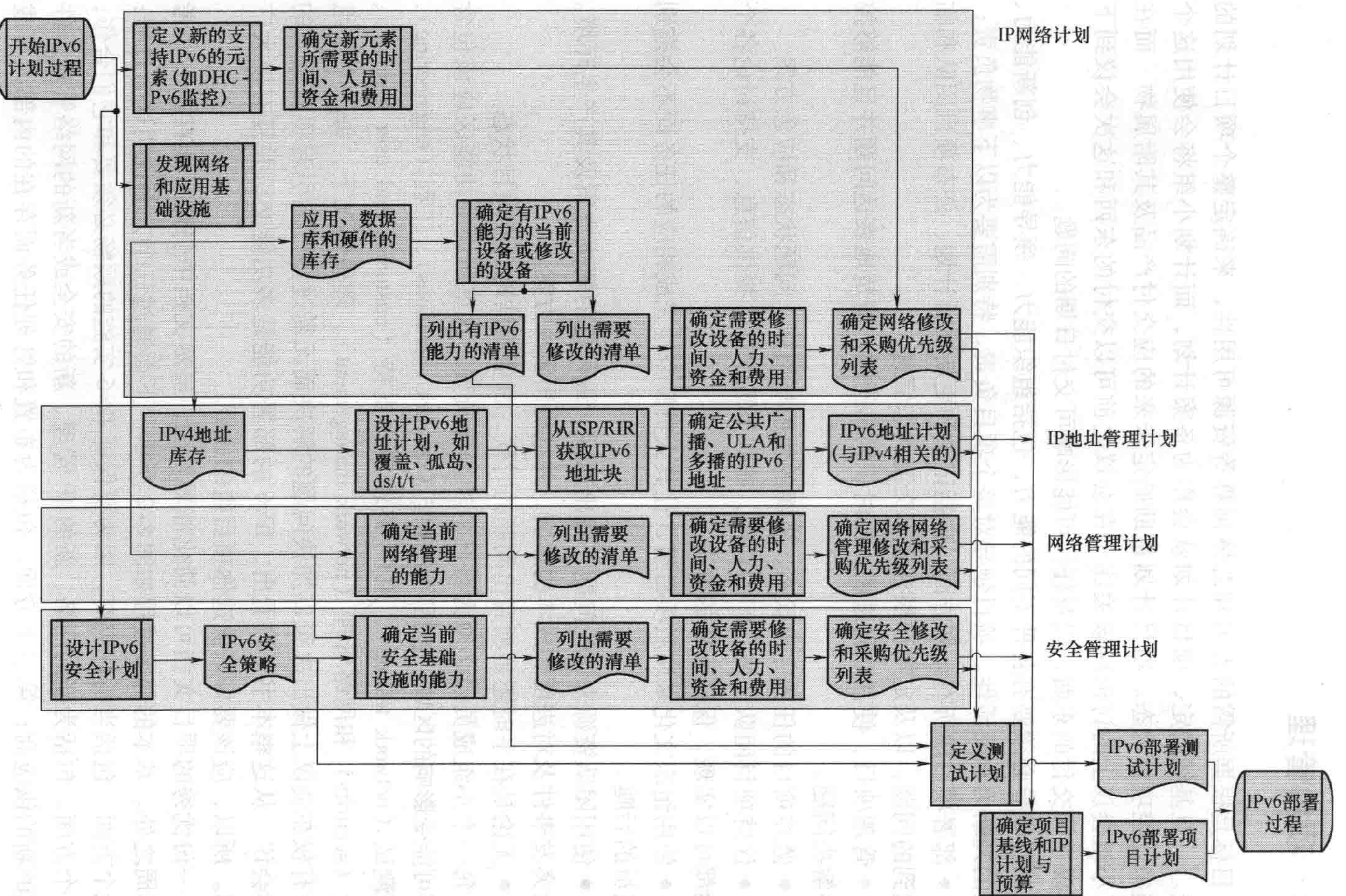



图 8-1 整体联合计划过程

## 8.2 项目管理

目标是根据所需的人力和总体预算等资源可用性，来确定整个项目计划的基线。一旦基线确定，项目计划就会作为备案计划，而计划小组将会使用这个文档来监控项目状态。项目计划帮助确定未来的应交付产品及其附属物。而在项目小组会议上的讨论应该关注于在会议之前可以交付的东西和这次会议到下次会议之间交付的东西，以及任何可能影响可交付日期的问题。

项目领导必须要有各种不同的能力，包括组织能力、领导能力、创新能力、良好的判断力和灵活性。项目领导对整个项目负责，特别需要为以下事物负责：

- 部署项目的所有计划方面，包括记录和更新项目计划、公布项目团队和临时遇到的问题，以及定期和根据需求交流状态信息。
- 管理项目，使其取得进展，同时识别出问题，清楚地表达问题并且能够控制并解决问题。
- 监控项目的任务完成状态、资源的使用与预算、问题状态和应急方案。
- 保持项目团队成员之间的合作，确保呈现状态、提供信息、支持讨论各个功能都可以实现，保证问题的解决。
- 使用组织文化来刺激项目组，让其人员勤奋地完成相应的任务而不会拖别的组员的后腿。
- 使用团队资源来讨论问题、找出原因、最可能的解决方案及其产生环境。讨论突发事件及对进度和成本的影响，并且得出解决方案。
- 无论是在平时还是项目出问题的时候，都要简洁地交流项目状态。

每一个小组成员还必须为自己的功能区负责。这里所说的功能区应该包含所有可能受影响的区域，包括网络操作（network operations）、工程（engineering），网络测试（network testing）、用户/终端客户支持（customer/end user support）、安全（security）和网络管理（network management）。理想情况下，每一个小组成员有权在会议上做出决策以加快问题的解决而无需过多可能引起项目延期的后续会议。从这整本书可以看出，IPv6 的部署功能是多功能交织并且有极大广度的。所以，应该要好好挑选你项目组的成员。

一份优秀的项目文档可以简化部署阶段，虽然文档中仍然可能存在一些考虑不周之处。基本的部署过程如图 8-2 所示。下面通过三个方面说明这个过程：第一个方面，包括当前供应商，因为要保证 IPv6 实施的顺畅必须要和它们合作；第二个方面，包括为新的网络、新的 IP 管理、新的安全或者新的网络管理组件寻找的新的供应商；第三个方面，包括为布置和管理任务而存在的内部人员或咨询人员。

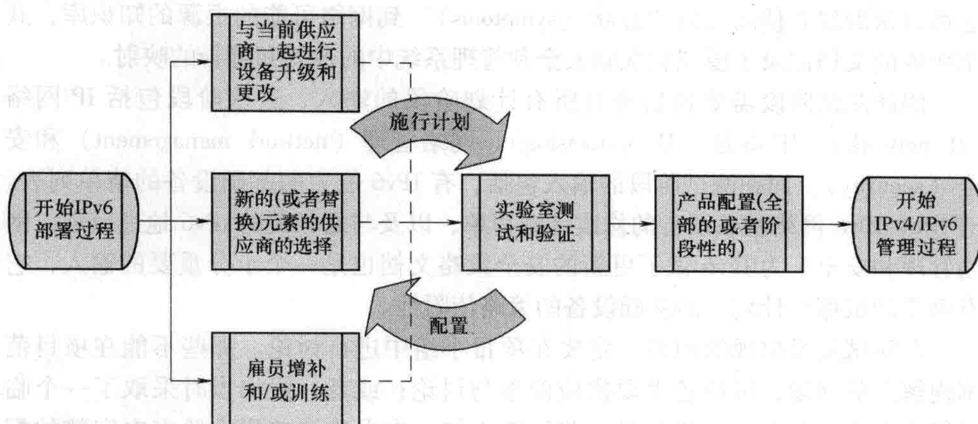



图8-2 基本的部署过程

如图8-2所示的迭代过程，定期的项目会议为“工作计划（working the plan）”提供了一个最初在实验室里处理和部署系统以便进行测试、确认进度和问题、反馈计划的修订版本和遇到突发情况的机会。当不可避免的问题出现的时候，每一个问题都被记录、讨论并尽快完美地解决是非常有必要的。这些解决方法的影响可能会涉及单个工作区到计划的主要变动，并且都会对资源和费用预算有一定影响。不过不管这个影响范围是怎样的，在解决方法中所有的问题都必须被监控。而项目计划也必须更新以反映任何资源变动（reflect any resource changes）、系统或基础设施替换（system or infrastructure substitutions）、进度表的超时和提前（schedule slips and advances）或者预算变动（budgetary changes）。

## 8.3 测试部署

正如之前所说的那样，IPv6的部署一定要先在一个非生产（non-production）环境或者一个实验室环境中被验证有效，以保证尽量少与IPv4生产网络（production IPv4 network）冲突。如果说项目计划定义了网络设备和网络系统的实施，那么测试计划就是这个系统在开始部署之前最后的关口。除了识别问题，测试阶段还能够丰富IPv4-IPv6在地址分配、配置基础设施、安全和管理系统上的经验，而且可以发现两个协议相互作用时候会出现的问题。除了IPv6待开发部分的部署之外，将IPv4和IPv6一起测试是值得推荐的。

网络中错误的情况和模拟的安全攻击都应该包括在实验室测试阶段，以便确认网络设备的响应，以及安全和管理系统所对应的检测通告。稍稍调整网络或者设备的参数设置并重新测试，可以帮助找到最好的参数。另外，文档和日

志可以帮助建立侦测到的“症状（symptoms）”到网络可能的根源的知识库，其中所说的文档记录了模拟状况到安全和管理系统中的通知和警告的映射。

因此测试阶段需要包括来自所有计划阶段的输入，这些阶段包括 IP 网络（IP network）、IP 寻址（IP addressing）、网络管理（network management）和安全（security）。而在测试阶段的输入包括，有 IPv6 能力的基础设备的清单列表、对 IPv4-IPv6 网络基础设备的修改计划清单，以及与之相关的 IPv6 地址计划、网络管理和安全。为 IPv6 做了更新的安全策略文档也是一个十分重要的输入，它有助于测试那些计划中的基础设备的策略依附性。

在测试阶段出现的问题一定要在项目小组中进行讨论。那些不能在项目范围内解决的问题，可能还需要供应商参与讨论；或者，假如当时采取了一个临时解决方案，那么还需要部置一些后期工作。本书建议使用实验室中完整的配置，来简化对升级、正在进行的故障排除及生产网络的问题重现（reproduction of production network problems）的测试。

## 8.4 生产管理

测试完成以后很可能会产生一个错误列表，这个列表会列出错误用例、检测到的异常和相关的网络影响。项目团队和高层领导需要确定是否有一些错误会阻止首次发布。如果是，这个错误必须马上解决，或者通过限定一个可接受的备选方案来解决。当这种阻碍被很好地处理以后，剩下的问题应该是在问题影响范围和所做修改方面上进行文档化，以及是否在生产网络中适当地监控事件。

在初始部署以后，生产网络应该马上被密切监控。实验室测试帮助描绘了 IPv6 的流量特征，但是现实生产环境中往往会出现一些意想不到的情况。所以最好密切监控 IPv6 的流量和安全日志。在分阶段部署中，在初始部署之后不那么需要实验室测试，除非有一个新的部件被添加到这个网络中。在实验室中得到一个比较满意的运行等级之后，随后就要密切监控的部署阶段，该阶段在生产环境计划范围内全面部署 IPv6 的各个部件。

# 第 9 章　管理 IPv4/IPv6 网络

基于良好管理计划的 IPv6 部署，为成功的部署和 IPv4/IPv6 网络部署后的持续管理建立了基础。IPv6 部署可能完全与 IPv4 网络重叠或者可能只共享网络的一个相同部分（如只是与因特网相连的部分），而 IPv4 仍然盛行（就目前来说）。另外，如果部署计划包括多个阶段，则每个阶段必须测试并且进入生产，包括合并到网络管理域中。如本书第 7 章介绍的，经过对网络管理的准备，现在是时候从这个配置、监控、维护和管理网络的谨慎计划中获得回报。

鉴于"网络"在 IPv6 部署之后仍然是那个"网络"，物理管理范围似乎变化不大，在程序上或许更是如此。从现在开始，在 M&P 方面必须考虑 IPv6 的地址计划，以进行 IPv6 连接的跟踪、安全和故障排除，以及与 IPv6 相关的统计监测和分析。本章将讨论这些问题。

## 9.1　常见的网络管理任务

回顾本书第 7 章介绍的 FCAPS 的基本功能分类，下面将从"配置"开始讨论一些常见的管理任务，然后再讨论其他类别。在描述这些任务时，已经遵循了所需的步骤。你的过程中可能还需要其他的文档和审批步骤。

## 9.2　配置管理

配置管理是一个基本的网络管理功能，并在初始部署后由技术或业务等举措驱动，如添加（或修改）一个新设备或新的网络服务，或创立一个新的分支机构。本节将讨论一些与管理 IP 地址空间和设备配置相关的常见且特别需要的任务。这些任务涉及日常由业务驱动的活动，影响着网络策划者对移动、添加和变更设备、子网及新的办公地点的考虑。该配置管理功能的结果是，每一个网络的核心要素，包括路由器、交换机、防火墙和像 DHCP 服务器和 DNS 这样的网络服务，都必须被配置以便在网络中各司其职。其他特定设备的参数通常需要被配置为与之前一样，可能会增加额外的 IPv6 配置信息。

### 9.2.1　网络中与配置相关的任务

下面来考虑这样一个例子，需要规划一个新的办公地点，可以是一个零售

商店、分支机构或是需要 IPv4 和 IPv6 支持新的网站。对网络规划者来说，配置管理涉及地址空间的规划、对现有和新的基础设施（静态的）的 IPv4 和 IPv6 地址的分配、其他设备地址的判定和分配、安全策略的实现，以及设备特定的网络信息配置（其中包括 DHCP/DHCPv6 服务器和 DNS 配置）。除了分配的 IPv4 和 IPv6 子网[⊖]外，如果地址空间已经根据每个应用程序进行了分割，则可能还需要多个子网，如本书第 5 章地址计划里提出的 VoIP 地址和无线地址。

假定已提供新网站地址分配的典型的模型或“模板”，试确定其版本（IPv4 与 IPv6）、类型（VoIP 与无线等），以及所需子网的大小。需要注意的是，鉴于/64的 IPv6 子网的纯容量，你可能选择在网站中为所有服务分配一个 IPv6 子网，以此简化你的子网划分计划和简单地配置多个 DHCPv6 池。对于每个子网，找出可用的子网地址，进而产生对给定的位置和应用的地址分配计划，以及 IP 地址规划“数据库”中的子网分配。这些是根据这个级别网络层次的配置策略进行的。比方说，用最合适的方法为每种类型分配下一个可用的子网。这就需要选择规模达到或超过每种类型所需大小的最小可用子网。

除了识别和记录每一个已分配的子网，子网分配的过程中，需要在适当的路由器接口配置每个子网地址和配置 IPv6 子网前缀选项，如在路由器通告中设置或清除“M”和“O”位以显示 DHCPv6 的可用性。当使用 DHCP/DHCPv6 时，在本地路由器中可能还需要对“helper”（中继代理）地址进行配置。一些子网上单独的 IP 地址需要分配给基础设备，如每个分配的子网路由器和服务器。定义和更新 DHCPv6 服务器的配置也需要设置地址池，以及相应的 DHCPv6 选项和（或）客户端类参数。这些都是在子网中使用 DHCPv6 的设备所需要的。

现在和未来，在子网分配有地址的设备可能需要 DNS 的名称解析信息。至少，这个信息适用于正向域的域名到 IP 地址查询，以及反向域的 IP 地址到名称查找。这需要通过域更新（如 in-addr. arpa 域和 ip6. arpa 域）和为名称服务器与静态分配的地址所做的资源记录更新，以达到定义和更新 DNS 配置的目的。当然，这些域必须存在或被调配和配置在各自的 DNS 上。

根据域的拓扑结构，可以利用现有的区域增加新的子网到一个位置，虽然这不是必须的。一个新域可能需要被定义或配置成一个子域或一个在适当 DNS 上的全新的区域。同样的，可能也需要添加与子网地址一致的反向域，除非更高层的 in-addr. arpa 或 ip6. arpa 区域拥有相应的 PTR 资源记录。

试考虑基于新分配的网络和设备对你的安全策略造成的影响。过滤 ACL 可能需要升级到能让子网上的设备穿过防火墙。DNS 的 ACL 可能也需要更新，以

---

⊖ 如果打算使用隧道或转换技术支持 IPv6，IPv6 子网配置可能不是必要的，尽管两者可能分别需要主机配置或转换网关配置。

适应新分配的子网上的服务器或设备的动态 DNS 更新。

子网分配过程说明了地址分配、指派、安全和路由器策略，以及 DHCPv6 服务器和 DNS 配置任务之间紧密的相互关系。根据你的业务流程，子网可以在地址指定和网络元素的配置之前被分配或保留。然而，以下这一整套步骤通常需要投入一个子网：

- 识别在需要每个子网的网络拓扑范围内可用的地址空间。
- 从相应的地址空间中为每个子网分配所需要的大小，同时记录 IP 地址计划中的分配情况。
- 更新关于已配置网络的路由配置。
- 分配和提供手动分配地址给路由器、服务器或其他的子网基础设备。
- 根据需要设计和配置 DHCP/DHCPv6 地址池来为子网上的动态主机提供服务。这可能需要选项、指令和客户类别的关联。该客户类别基于为地址池的使用而定制的设备的需求。
- 定义服务子网上的主机所需要的新的 DNS 域，在新的或已存在的域中为基础设施或静态设备定义资源记录，并配置合适的 DNS[⊖]。
- 考虑允许或限制每个已分配子网之间的可达性所需要的安全策略更新。
- 通过确认子网的配置和可达性，以及网络合适度和应用性能，此外也通过验证相应的 DHCP 和 DNS 配置和其他一些核心网络服务（如 NTP），来完成分配过程。

## 9.2.2 添加新设备

向网络中添加一个新的设备，涉及配置其 IP 地址或分配方法，配置设备上的安全设置，并设置特定于应用程序的参数。分配、取消分配和重新分配 IP 地址到单个主机往往是在许多组织中最常见的配置管理活动。这通常类似设备的部署、重新部署或拆除，包括路由器、服务器、打印机等。就地址分配而言，IP 地址清单数据库应满足通过查询，立刻确定可用的 IP 地址。虽然把整个库存保证作为一个单独的任务来讨论，但是只是为了验证库存的准确性而去 ping 每个被分配的候选 IP 可能有用的。被分配的 IPv4 和/或 IPv6 地址，应指派给网络清单数据库中指定的设备。如果正在使用在这个特定的子网上支持的 IPv6 主机端点隧道，请注意清单内相关的隧道地址。

物理 IP 地址可以通过 SLAAC 或通过使用 DHCP/DHCPv6，以手动（静态）配置设备的方式进行。在静态分配的情况下，分配的地址必须直接在设备上进

---

⊖ 在某些网络中，DHCP 地址的资源记录的预配置是必需的，以允许这些地址的用户在没有执行动态 DNS 更新的情况下出现在 DNS 中（如方便 VPN 的连接，该连接需要 PTR 记录的存在）。

行配置，因此除非 IP 地址分配者也同时负责物理分配，这个过程可能需要一个电子邮件或电话来把输入的 IP 地址信息传达给设备所有者。如果启用 SLAAC，该设备将自动配置 IPv6 地址；确定选择什么地址将要求本地的“控制台”能够显示它或轮询路由器的邻节点表。当使用 DHCP 时，如果需要一个确定的 IPv6 地址，一条在适当的 DHCP 服务器配置文件中的条目可能对设备的 DUID 到分配的 IP 地址的映射十分重要。

大多数具有 IP 地址的设备需要相应的 DNS 资源记录使得通过主机名实现可达。使用地址分配的 DHCP 方法，DHCP 服务器可以被配置来更新用于 IP 地址分配的主 DNS。此更新将影响用于域名到 IP 地址（A/AAAA）查找的正向域和用于反向（PTR）查询反向域。如果手动分配地址，需要一个类似的 DNS 更新的任务。用新的主机信息更新 DNS 可能需要编辑或更新服务器上相应的区域文件，或通过发送动态更新。

你可能不想要自动配置的设备靠自己更新 DNS，至少在一个企业网络中是如此，虽然这可能适合于社区或特定的网络。识别一个新的自动配置的设备的存在，以手动更新 DNS，具有一定的挑战！如果此类设备需要在 DNS 中的解析信息，可能需要使用路由器日志或子网窥探工具来确定 IPv6 地址。

根据所添加设备的类型，在设备上配置的安全策略可能包括定义 ACL，配置主机防火墙和（或）初始化其他预防性安全措施，如反恶意软件的软件。如果需要，风险管理系统可能需要进行初始化，以使设备的配置和管理得以进行。

总之，增加一个新设备的任务包括以下子任务：

- 确定设备将如何获得 IP 地址，通过手动配置、SLAAC 或 DHCP。
- 如果是 DHCP，如有地址池，确定子网上当前的地址池是否有支持设备的能力；如没有地址池，在 DHCP 服务器上配置对应 DHCP 类型的地址池及必要的选项参数。
- 如果是手动 DHCP（即该设备通过 DHCP 分配到一个“静态”的地址），在设备所在的子网内找出一个空闲的 IP 地址，并通过配置 DHCP 服务器将该地址分配给设备，从而为设备的 DUID 设置保留或分配手动的 DHCP 地址。
- 如果是在设备上手动配置，在设备所在的子网中找出一个空闲的 IP 地址，并将该地址分配给设备。将所分配的静态 IP 地址手动配置在设备上。如果可能的话，选择一个非单调增加的（即随机的）接口 ID。
- 在所有情况下，用已分配的地址来更新 IP 地址计划，无论是使用电子表格或使用其他 IPAM 工具。
- 确定是否需要手动创建和更新 DNS 资源记录。对静态分配地址来说，通常是如此。对于 DHCP 分配的设备，DHCP 服务器可以被配置来执行动态更新。虽然在动态更新对于政策不可行或不被政策允许的情况下，可能需要相应资源

记录的手动更新。

- 配置设备的安全软件和策略。
- 如有必要，配置设备管理参数。
- 通过 ping 成功和验证其 DNS 中的资源记录，来验证地址分配过程是否完成。通过地址池分配的地址的设备，可能不需要验证。然而，如果需要，则该地址可能不是事先已知的。在这种情况下，定位设备的 MAC 地址或 DHCP 租约文件中的 DUID，然后由相应地址的 ping 证实其是否被成功分配。

### 9.2.3 删除任务

地址分配是一个自上而下的过程，从你的基础部署分层块的分配，然后子网可被分配，IP 地址可被分配。删除设备或地址空间需要进行逆操作，且必须是自底向上的。在底层的块、子网和设备删除之前，删除一个地址块将使这些底层元素陷于困境。

#### 9.2.3.1 删除设备

从网络中删除设备的操作相对简单：根据需要，在网络库存和相应的安全和管理系统中将设备表示为被删除或待删除对象。这包括从 DHCP 服务器上移除任何 M-DHCP 条目（如果合适的话），释放租赁，并删除相关的 DNS 资源记录。但是，在把地址分配给另一台设备之前，必须小心确保该地址已被设备放弃且 DHCP 和 DNS 更新已完成。例如，仅删除一个 DHCP 服务器上的租赁并不会迫使持有该租赁的客户放弃它。DHCP 配置消息旨在迫使 DHCP 客户进入更新状态，使一台服务器可能确认响应客户更新租约的请求，从而释放地址。然而，这还没有被广泛实现。

将地址表示为“待删除”或类似的状态，会提醒其他管理员不要将该地址分配给另一个设备，直到收到对其可用性的确认。这个确认过程需要 ping 地址，也许先后超过数天，并确认在 DNS 和 DHCP 服务器中与其相关联的数据被删除。

如果被删除的设备本身是网络元件，则需要关注被删除 IP 地址的 ACL 或相关的信息。例如，如果删除一台 DNS，其地址应从相应的“允许更新的”或有关的 DNS ACL 中删除。

#### 9.2.3.2 删除子网

在关闭一个网站或合并地址空间时，可能需要删除一个子网。在被删除的子网中带 IP 地址的设备应该被移除或退役，这样的子网才是没有地址分配（或许不同于其他子网服务的路由器）。在所有 IP 地址已确认为空闲的之后，子网可能被回收到空闲地址空间用于未来的分配。

在释放一个子网时，可能会将释放的空间合并到一个连续的地址块中，从而得到一个更大的空闲块用于未来的分配。与防火墙和 DNS ACL 配置相关的其

他日常任务应考虑被删除子网、域、ACL 和资源记录。

### 9.2.4 地址重编或移动任务

移动或重编地址块、子网或独立地址基本上结合了分配过程和删除过程。如上所述，分配过程应以自上而下的方式，对那些底层子网和 IP 地址将被移动到的空间进行分配。当地址被移动到分配的目标空间后，删除过程从下往上地释放了地址空间。在本质上，分配给被移动的地址范围的大小必须能容纳被移动地址，临时地址空间应为与设备组相关联的地址空间的两倍。当地址被移动，前者的地址空间可被释放，返回地址分配到以前的水平。

#### 9.2.4.1 设备移动

从 IP 地址分配的角度来看，移动服务器或其他设备可看作在目标子网分配 IPv4/IPv6 地址和在移动完成后删除当前子网的 IP 地址两者的结合。根据地址分配的方法和移动的类型，可以使用不同的策略。移动类型包括设备物理移动到不同的子网（物理移动）和在同一或不同的子网里重新分配的 IP 地址（逻辑移动或重编）。非移动 IP 设备的物理移动，如服务器，通常会涉及关闭电源和重新启动，提供了更多的地址分配过程中的时机控制。

*物理移动* 物理移动意味着将设备断电、移动，然后在目标位置通电。这里假定在其他网络设备移动之前，路由器和交换机已经在目标位置完成了安装和配置。这种方法有助于最大限度地减少停机时间，否则将体验一个完整的“拿起然后移动”的方式。对于 DHCP 和 DHCPv6 分配的设备，如果整个池被移动，应在 DHCP 服务器上建立与每种所需 IP 地址类型对应的目标池。确保服务目标子网的路由器被配置为中继 DHCP 报文到与新池配置在一起的 DHCP 服务器。你还可以为 DHCPv6 配置中继，尽管知名多播地址已被定义回避这一要求。

当被移动设备在新位置启动，它们可能会尝试更新最近它们在原子网上持有的租约。DHCPv6 服务器会对每个客户的请求消息发出回复，表明不能更新有关租约。客户端将重新初始化并发出一个请求数据报，以取得新的租约。DHCPv6服务器从新的目的池中发放 IPv6 地址租约。一旦所有设备完成物理移动，服务于原子网的池可能会退役。

手动 DHCP（Manual DHCP，M-DHCP）设备就像是一个自举协议（bootp）设备，每次与 DHCP 交换时都接收相同的 IP 地址。这需要在 DHCPv6 服务器的配置中有一个预配置的一对一的关联关系，即设备的 DUID 与一个固定的 IPv6 地址对应。一个 M-DHCP 设备的物理移动需要在服务于新子网的 DHCPv6 服务器中创建 M-DHCP（“主机”）条目，且在之后删除在原 DHCPv6 服务器上的条目。如果正在使用相同的 DHCPv6 服务器，则只需编辑与设备的 DUID 相关的 IP 地址。设备在新子网上电时，它应该遵循上述类似的过程，伴随着 DHCPv6 过程

的重新初始化。

移动一个自动配置 IPv6 地址的设备，将导致设备开机后会通过路由器发现和相应的子网政策（包括 DHCPv6 服务的可用性）来检测其新的子网。设备如果使用 SLAAC，则自动配置其 IPv6 地址，然后通过重复地址检测验证其唯一性。如果使用 DHCPv6，伴随的是正常的 DHCPv6 过程，用以获取 IPv6 地址和相关参数。在某些情况下（即当路由器通告中 O 位被设置且 M 未设置），自动配置和 DHCPv6 均可使用。

DNS 资源记录的更新可能由 DHCP 服务器在分配地址时进行。这个过程可以更新正向（A/AAAA）和反向区域记录（即 PTR）。如果 SLAAC 被支持且 DNS 更新可以信任，终端客户端可能会被配置为自动更新 DNS。该信任模型是依赖组织的。一些组织放弃 DNS 更新，因为考虑到这种主机将无法通过名称访问。另一种选择是使用 IPAM 系统或使用动态主机的 A/AAAA 记录和 PTR 记录手动更新 DNS。

手动配置的设备的物理移动，需要从 IP 清单库中分配一个地址，并当它在新子网启动后手动配置新 IP 地址。这时，旧的地址可以在库存中释放出来，尽管临时的“待删除”状态可能对防止相应的地址在核实可用性前被过早地重新分配是有用的。DNS 资源记录也应该被更新；以反映设备的新的 IP 地址。

在所有这些情况下，IP 库存应被用来识别可用的目标子网或地址池的可用地址和释放原子网上的地址，以及设备的移动被确认后相应的 DNS 资源记录。任何用以追踪 IP 地址的设备库存系统同样应该被更新。

任何影响被移动设备的 ACL 或基于地址的过滤规则，也需要被更新；还有任何与地址相关的网络管理系统的轮询配置需要被更新。同样，任何“硬编码” IP 地址受移动影响的配置文件或应用程序也必须被更新。

*逻辑移动* 逻辑移动更具挑战性一些，因为它们不一定涉及设备的重新初始化。对于 DHCP 设备，包含所需目的 IP 地址的地址池应在（相同或不同的）DHCP 服务器上进行配置。当前地址池的租约时间应在移动的时间之前结束。例如，如果正常的租约时间为 1 周，它应该被降低到的设备移动的一周内的 1 天和设备移动的一天内的 2 ~ 6 小时。一台设备可能刚好在你将租约时间变为天之前续签了一份为期 1 周的租约，所以该租约在半周以内不会尝试更新（或基于你的 DHCP T1 时间选项设置）。因此，如果你名义上的租约时间为 2 周，请在计划移动前的最多 2 周时，缩短租约时间。在移动的当天，如有必要，把租约时间设置为最短的[⊖]时间，这样所有设备将几乎在同一时间移动。如果同时移动不是

⊖ 最短时间可以是几分钟或几小时，这取决于网络流量和服务器性能方面的考虑。租约时间越短，将发送更多的 DHCP 报文，但也使 DHCP 客户端的移动时序协调更一致。

关键的，让租约时间分布在几个小时上，则在几个小时内应该能产生一个完整的移动。

在这种情况下，如果可能，建议由 DHCP 服务器执行 DNS 更新，以便更为接近地通过地址的变更映射出 DNS 信息的更新。手动干预一个“无限租用的”IPv6 设备可能是必要的，除非它在拥有无限的租约后仍符合租约更新政策。另一种选择是，如果客户端支持重新配置。则命令 DHCPv6 服务器发送重新配置信号给客户端，指示它重新开始 DHCPv6 的过程，

手动分配地址的设备的移动与物理移动遵循相同的过程。一组目的 IP 地址从 IP 库存中被分配出来，且每个新的 IP 地址都配置在设备上。一旦经过确认，每个旧地址可以被分别释放。DNS 资源记录也应更新，以反映该设备的新的 IP 地址。

自动配置设备的逻辑移动，可以通过配置服务于相应子网的路由器来执行，以降低它在相邻路由器发现过程中宣布的优先和有效的地址生存期值。当用一个“正常”的地址引入新的前缀，缩短这些来自待移动设备的地址前缀的定时器值，则生存期将使自动配置的设备能够自动执行此逻辑移动。当所有装置被移动且原前缀的有效生命期期满时，该前缀可以被删除。

#### 9.2.4.2 子网移动

移动子网可能涉及两种结果之一：子网和已分配的 IP 地址移动到另一个路由器接口；保存当前的地址分配或移动到另一个路由器或接口，需要新的子网地址。在后一种情况下会包含子网重编任务，因其也导致了一个新的子网地址，虽然不必将子网移动到另一个路由器接口。第一种情况需要考虑地址空间在层次上汇总，但大体由路由器配置的修改和验证组成，这些配置符合地址计划、受影响的安全和过滤政策的更新和网络监控系统，以及对路由表和 DHCP 中继地址的更新（如有必要）。

由于物理移动或更高层次的重编产生的子网移动，通常需要多一点的工作量。设备被真实移动的物理移动（如当一个办公室被移动），是具有本质破坏性的。目标子网可能被分配和配置在目标路由器的接口上，连同上述与预留静态地址，更新 DHCP/DHCPv6、DNS、防火墙和网络管理配置相关的其他任务。每个被移动的设备接入时，需要手动分配新的地址和（或）从地址池上获得一个上述用于 IP 地址移动的与子网相关的 DHCP/DHCPv6 租约。自动配置的设备将检测到它们通过 NDP 正在连接的子网的前缀，并应相应地为每个前缀自动配置一个单播 IPv6 地址。逻辑子网移动或重编同样遵循设备的逻辑 IP 地址移动的过程。

在所有的设备已从旧子网移动到新子网后，原子网地址空间可能会伴随着子网删除过程被释放出来。

### 9.2.5 块/子网分割

分割地址块需要从给定的源块中创建两个或多个更小的块。分割对地址空间的释放可能是必要的，甚至可以作为一种地址空间子分配的方法。在前一种情况下，一个子网内的地址可能会被合并到子网前半部分，同时释放后半部分子网的分配。这样，块分裂将产生一个被占有子网（前半部分）和一个空闲子网（后半部分）。一些组织拥有已分配的区域块，然后将它们分割后分配给在较低地址层的子块和子网。从某种意义上说，这是块分配的一种形式。

要认识到 DNS 反向区域分裂块时会产生影响。如果两个子网的 DNS 行政机构在一组管理员下保持统一，原来的正向或反向区域可能不需要修改。然而，如果得到的分割块或子网让它的设备在 DNS 中由不同的授权的机构管理，则原来的反向域将同样需要分割。这需要建立两个与产生的分裂子网相对应的反向区域，并通知被分割的父反向区域的管理者，其负责将反向区域树妥善委托给合适的用于权威信息的 DNS 组。

分割块不需要被限制为仅在对半分裂，把一个/60 分成两个/61。拆分可以用来从一个/60 中开拓出一个/64，虽然一般来说，使用的子网分配过程来执行也许会更好。这样的分裂会产生所需的/64，以及由一个/64、/63、/62 和/61 组成的自由空间。在这个例子中，要遵循最优分配策略保留大的块。另外，可以直接将一个/60 分裂成 16 个/64，其中包含了统一分配的政策，而不是需要时“最适合”的分配策略。

总之，分割块的过程与分配块相似。被分割的块被依次划分，直到获得所需的块大小。剩余的空闲块或者被保留，或者也被分割至所需块的大小，以呈现统一的块分割。DNS 关于反向区域树和行政委托的含义必须加以考虑。还有，请记住每个分割产生的网络产生一个额外的网络地址和广播地址。

### 9.2.6 块/子网连接

将两个相邻的同样大小的地址块或子网组合连接成一个单一的块或子网。例如，块 2001:DB8:0:2::/64 和块 2001:DB8:0:3::/64 可以连接形成块 2001:DB8:0:2::/63，因为两个/64 共享同一个 63 位的前缀。请注意，如果试图连接块 2001:DB8:0:1::/64 和块 2001:DB8:0:2::/64，情况则并非如此，因为这不是一个有效的连接。为什么想要连接？因为连接能使较大的块聚合，可供未来分配。稀疏分配使得增加地址空间容量，而无需增长路由表。在本书第 5 章关于稀疏分配的讨论中，可以看到一个这样的例子。分配可能请求任意大小的块或子网，因此聚合块空间也增加了在需要时可用与分配更大的块可能，这反过来又最大限度地减少在路由协议中需要公告的网络数量。

汇总连接块，可能还需要一个 DNS 反向区域的更新，以使基础设备资源记录合并成“连接的”反向区域反映出生成的统一子网和安全过滤策略的更新。

### 9.2.7 DHCPv6 服务器配置

DHCPv6 服务器配置是一个关键的地址管理任务，胜于地址池的创建、移动和删除，虽然额外功能的范围受 DHCPv6 服务器供应商的能力限制的。DHCPv6 服务器配置参数中关键项如下：

- DHCPv6 的前缀池，负责用于分配给请求前缀委派的路由器的 IPv6 块组。
- DHCPv6 地址池，负责地址范围和相关 DHCPv6 选项，以及服务于动态、自动和手动的基于 DUID 的 DHCPv6 客户端的服务器政策。
- 客户端类，负责参数的匹配值（如供应商类标识符为“Avaya4600”）、相关的允许/拒绝池、DHCPv6 选项和服务器策略。
- 分割范围的高可用性参数设置（DHCPv6 的故障转移目前 IETF 内部正在研究）。
- 服务器活动的配置，如动态 DNS 更新和其他的服务器指令与参数。

实际服务器的配置语法和接口将取决于服务器类型。例如，ISC DHCPv6 服务器可以通过编辑 dhcp. conf 文件来配置，而微软 DHCP 可以通过使用 Windows MMC 接口进行更新。这两者和其他 DHCP 供应商也提供命令行接口或 API 来执行配置更新。关于此类和其他产品，请查阅你的供应商的文档。

个别静态 IPv6 地址分配需要进行记录，以确保其唯一性。在分配的子网内，应对 DHCPv6 地址池进行跟踪以提供一个子网内地址分配的整体视图，不管是静态还是动态分配的。当在电子表格内对独立 DHCPv6 租约的跟踪未能容易地执行时，应该至少执行对电子表格或数据库内的地址池的记录。随着时间的推移，这将有助于确保唯一地址的分配。

这样的跟踪也是有必要的，以便于关联给定主机的多个地址，如双协议栈。这种综合的地址分配数据存储将提供 IP 地址的库存的已知程度。图 9-1 给出了 IP 地址的示例子网库存表。

请注意图中对 IPv4 地址为 10.17.2.5 的保留设备的双栈地址分配。IPv6 地址由子网前缀和一个直观地映射到相应的 IPv4 地址的接口 ID 连接在一起的接口 ID 连接组成。接口 ID 并不是 IPv4 地址的二进制映像，而是一个“可视化的”映像，以便于 IPv4 地址能够通过视觉容易地区分出来。这种地址分配方法无疑帮助了日志或管理系统中检测到的设备地址映射到知名或熟悉的 IPv4 地址，但同时也暴露了一个潜在的安全漏洞。该漏洞涉及允许攻击者通过防御能力可能更弱的 IPv6 地址锁定一个给定的已知 IPv4 地址的设备。

| 欧洲-西方-罗马-VoIP子网 | | | | | |
|---|---|---|---|---|---|
| 10.17.2.0/24 | 2001:db8:4af0:8812::/64 | | | | |
| | | | | | |
| IPv4地址 | IPv6地址 | 主机名 | 设备类型 | 分配方法 | 硬件地址 |
| 10.17.2.1 | 2001:db8:4af0:8812:ca00:21ff:fe07:39f1 | roma-core01 | 网关/路由器 | Manual | C8-00-21-07-39-F1 |
| 10.17.2.2 | 2001:db8:4af0:8812:ca00:22ff:fefe:a901 | roma-core02 | 网关/路由器 | Manual | C8-00-22-FE-A9-01 |
| 10.17.2.3 | 2001:db8:4af0:8812:212:65ff:fe91:27 | roma-ops01 | 交换机 | Manual | 00-12-65-91-00-27 |
| 10.17.2.4 | 2001:db8:4af0:8812:212:65ff:fe91:1ab1 | roma-ops02 | 交换机 | Manual | 00-12-65-91-1E-B1 |
| 10.17.2.5 | 2001:db8:4af0:8812:10:17:2:5 | | | Reserved | |
| 10.17.2.6 | 2001:db8:4af0:8812:10:17:2:6 | | | Reserved | |
| 10.17.2.7 | 2001:db8:4af0:8812:10:17:2:7 | | | Reserved | |
| 10.17.2.8 | 2001:db8:4af0:8812:476a:1ff:fe00:d98 | romops-print01 | 打印机 | M-DHCP | 45-6A-01-00-0D-98 |
| 10.17.2.9 | 2001:db8:4af0:8812:476a:1ff:fe20:3df0 | romops-print02 | 打印机 | M-DHCP | 45-6A-01-20-3D-F0 |
| 10.17.2.10 | 2001:db8:4af0:8812:476a:1ff:fe01:65d1 | romops-print03 | 打印机 | M-DHCP | 45-6A-01-01-65-D1 |
| 10.17.2.11 | 2001:db8:4af0:8812:476a:1ff:fe94:309e | romops-print04 | 打印机 | M-DHCP | 45-6A-01-94-30-9E |
| 10.17.2.12 | 2001:db8:4af0:8812:476a:1ff:fe89:a20c | romops-print05 | 打印机 | M-DHCP | 45-6A-01-89-A2-0C |
| 10.17.2.13 | 2001:db8:4af0:8812:476a:1ff:fe0a:a98b | romops-print06 | 打印机 | M-DHCP | 45-6A-01-0A-A9-8B |
| 10.17.2.14 | 2001:db8:4af0:8812:476a:1ff:fe49:1fe | romops-print07 | 打印机 | M-DHCP | 45-6A-01-49-01-FE |
| 10.17.2.15 | 2001:db8:4af0:8812:10d::78f3 | opsfile41 | 服务器 | Manual | |
| 10.17.2.16 | 2001:db8:4af0:8812:8021:776:d:1ce1 | opsfile42 | 服务器 | Manual | |
| 10.17.2.17 | 2001:db8:4af0:8812:1107:ca1:50ee:7b | opsfile43 | 服务器 | Manual | |
| 10.17.2.18 | 2001:db8:4af0:8812:8e:320:dec1:90a:a8 | opsfile44 | 服务器 | Manual | |
| 10.17.2.19 | 2001:db8:4af0:8812::8e10:d878 | opsfile45 | 服务器 | Manual | |
| 10.17.2.20 | 2001:db8:4af0:8812:4:7dad:910:21a2 | opsfile46 | 服务器 | Manual | |
| 10.17.2.21-10.17.2.50 | | | | Reserved for net-servers | |
| 10.17.2.51-10.17.2.254 | | | VoIP 电话 | D-DHCP | |
| | 2001:db8:4af0:8812:ffff::/80 | | | DHCPv6 pool | |

图 9-1 IP 地址的示例子网库存表

## 9.2.8 DNS 配置

如 DHCP 服务器，DNS 的配置是一个重要的网络管理功能并与地址分配、指派、移动和删除紧紧地联系在一起。这些前面所讨论的任务影响着 DNS 域、资源记录，可能还有服务器的配置参数。DNS 配置参数的关键项如下：

- 域名，在 DNS 上添加、修改或删除域/区域。
- 资源记录，添加、修改或删除资源记录，如特别的 AAAA 和 PTR 类型。
- 服务器、视图和区域配置，设置和修改选项参数、影响着 ACL、服务器配置、DNS64 配置等。

实际 DNS 的配置的语法将取决于服务器类型。ISC BIND 服务器可以通过编辑服务器上的 named. conf 文件和相关的区域文件进行配置。支持 DDNS 的 DNS，可能也以这种方式支持资源记录的更新。对 nsupdate 或类似的 DDNS 机制的使用，提供了一种手段来执行增加的更新，而不必手动编辑区域文本文件和重新加载各自的区域，如对 rndc 的使用。DDNS 更新只适用于资源记录的添加、更改和删除，因此任何区域或服务器配置参数的变更或区域的增删，仍然需要文本文件的编辑，重载 named. conf 文件和（或）受影响的区域。

考虑到 IP 地址与反向域、主机名和其他主机的信息之间的直接关系，以及 DNS 能通过名称导航而不是 IPv6 地址，所以显然 DNS 是管理 IPv6 网络的一个关键因素。DNS 提供了主机名和 IP 地址之间的重要联系，使得 IP 应用的使用更为容易。

从 IP 地址管理的角度来看，显然反向 DNS 域与 IP 地址块和子网的分配有直接关联。这些域直接来自其对应的 IP 地址。

为了与集中的 IP 地址库存的理念保持一致，应对与每个 IP 地址相关的主机名和资源记录进行跟踪。参考图 9-1 所示的 IP 清单电子表格，如果意大利罗马办事处管理其自己的 DNS 区域，其管理员需要为 rome. ipamworldwide. com（前向域）2. 17. 10. in- addr. arpa 和 2. 1. 8. 8. 0. f. a. 4. 8. b. d. 0. 1. 0. 0. 2. ip6. arpa，管理区域文件。在这些区域文件中，资源记录条目需要维护。例如，对于 IPv4 地址为 10. 17. 2. 15 的主机 opsfile41，必须生成以下区域文件条目：

- rome. ipamworldwied. com 的区域

| | | | |
|---|---|---|---|
| opsfile41 | IN | A | 10. 17. 2. 15 |
| opsfile41 | IN | AAAA | 2001：db8：4af0：8812：10d：：78f3 |

- 2. 17. 10. in- addr. arpa 的区域

| | | | |
|---|---|---|---|
| 15 | IN | PTR | opsfile41. rome. ipamworldwide. com. |

- 2. 1. 8. 8. 0. f. a. 4. 8. b. d. 0. 1. 0. 0. 2. ip6. arpa 的区域

| | | | |
|---|---|---|---|
| 3. f. 8. 7. 0. 0. 0. 0. 0. 0. 0. 0 d. 0. 1. 0 | IN | PTR | opsfile41. rome. ipamworldwide. com. |

需要确保把这个库存信息正确抄写到 DNS 配置中。从这个“数据库”，可以得出对应到每个主机的 A/AAAA 记录和 PTR 记录。可以扩大在电子表格中的列跟踪与给定主机关联的额外资源，如 CNAME，MX 记录等。在这个例子中，正在为静态定义的主机输入DNS 信息，而这个过程应该在子网的每台这样的主机上重复进行。从 DHCP/DHCPv6 地址池获得租约的主机或自动配置设备可以通过动态 DNS 更新其主机名信息。

### 9.2.9 前缀重编

服务供应商在你的网络中提供了按供应商分类汇总的 IPv6 地址空间供使用，在服务供应商发生变化的情况下，IPv6 前缀必然会发生变化。RFC 4192（即本书参考文献［99］）描述了IPv6 网络重新编号的过程。为了最大限度地减少中断，当即将被弃用的旧前缀仍是可操作的时候，就需要部署新的前缀。这就需要把新前缀依次子网化到子网层，使单个设备能够自动配置，或能够在移除旧前缀和保留连接之前，从新前缀中手动分配 IPv6 地址。这个过程类似最初在 IPv4 网络上部署 IPv6。相关步骤包括相应地根据子网信息配置路由器接口、传播新前缀的路由更新、配置安全过滤器和 ACL 以适应新的前缀，以及配置各自的反向区域的 DNS。此外，在应用程序、配置文件、DHCP 选项的参数值和其他

任何地方中的 IPv6 地址，各自都需要更新，以反映新的前缀。一旦新的前缀已被部署，网络有效地利用（至少）两个前缀。在移动子网和设备时，旧前缀上的地址应该随着时间的推移衰减，如通过缩短 DHCP 租约时间和减少公告的路由器首选和有效的寿命值。随着设备转变到新的前缀，旧前缀的地址变成空闲了，然后释放子网，再到块，最终上升到整个前缀本身。

## 9.3　故障管理

故障管理不仅包括故障检测，还有告警通知、故障隔离功能，故障跟踪和解决问题的过程。对 IPv6 网络元素和服务器针对故障与事件进行监测，能够以一种积极主动的手段最大限度地减少服务中断，并作为当前 IPv4 监测过程的一个扩展。正如本书第 7 章中讨论的，IPv6 的理想需求是目前监测的扩展、支持 IPv6 传输的故障管理工具、SNMP MIB、日志信息和相关的健康和状态数据，而不是新系统的使用。

### 9.3.1　故障监测

根据部署的网络元素和服务器所支持的功能不同，故障检测也可以使用各种方法进行。这些方法包括专有的轮询或通知、系统日志扫描和（或）转发、SNMP 轮询，以及基于 SNMP 的网络管理系统的陷阱检测。

除了监测由设备报告的网络设备的状态，它还对监测这些设备提供的服务十分有用，特别是关键的 DHCP、DNS、NTP 和其他的一些服务。监测该服务只需要发送相应协议的消息，并在合理的响应时间内收到并确认一个正确的响应。例如，定期向服务器发出 DNS 查询，将能检测 DNS 的服务是否正在运行，并且能够履行其解析 DNS 查询的职责。

监控网络设备和通信链路是普遍网络监控的一种常见的做法，而且能洞察影响客户访问核心网络服务器的能力的中断。这条附加的信息对解决某个具体问题或故障是非常有用的。故障相关性是从多个网络元素或管理系统收到的各个故障的分析，以帮助找出一组故障的根本原因。例如，对来自第 2 层交换机、路由器和广域网接入设备的故障进行集中分析可知，这三种故障是相关的，而且最可能的根本原因是链路中断。

故障相关性是大型网络管理系统的一个共同特点。无论故障相关是由网络管理系统自动执行或手动通过比较来自多个系统的信息，这个过程将暴露出更为广泛的故障分析的数据集，进而为给定的服务器、链接或网络元素隔离故障。作为一个双栈设备，从一个单独的设备的多个 IP 地址的轮询响应的简单故障相关性分析，可以帮助识别潜在的协议和路由问题。

### 9.3.2 故障排除和故障解析

IPv6 引入了一个用于故障排除的附加层，这是一个双刃剑。这为 IPv6 相关问题的识别和解决增加了一层复杂性，但同时也为数据收集和诊断提供了一个接入到设备的辅助协议。

## 9.4 计费管理

计费管理的基本目的是为了保持每个人都诚实。那些已分配的地址仍然在使用吗？有没有哪些没被分配的地址正在使用？新的子网在路由器上是否进行配置了？对于这类问题，计费管理要能够验证成功的配置，以及对 IP 网络和地址计划的整体遵守。计费管理功能的技术包括 IP 地址、路由器子网、交换机端口映射、设备接口信息、DNS 资源记录和 DHCP 租约文件的发掘。

为了把这些信息与 IP 清单的“记录计划进行比较”，分析发现的信息是必要的。这种差异的报告和比较是比较困难的工作，但为库存的准确性提供了一定程度的保证。如果没有这样的功能，恶意用户可以访问免费服务或以其他方式渗透网络。此外，尚未实现的计划网络的变化可能会导致下游过程的延误和内部或外部 SLA 的配置时间间隔上的冲突。

### 9.4.1 库存保证

到目前为止，已经介绍的每个常见的网络管理任务都依赖准确的设备清单，以使子网、设备、IP 地址及相应的路由与过滤配置的分配、删除和移动得以实现。准确性，对于这些地址管理任务是绝对必要的。但准确的库存对一般的故障排除也是必不可少的。一个远程站点是否应该由于网络中断而设置为不可达？这可能有必要为站点上的设备识别 IP 地址、资产信息或其他与网络相关的数据。在可能最需要这样的信息或当它不能直接从网络获得的时候，只有维护一个准确的网络清单时，这样的信息才能被访问。

本节将审查你可以采取的步骤，以确保你的网络库存的准确性。这包括控制谁可以对特定的 IP 和设备信息、实际网络的发现、为库存核对实际情况及地址空间的最终回收做出特定的改变。

#### 9.4.1.1 变更控制和管理员的责任

正如在审查这些网络管理任务时所看到的，网络和 IP 地址库存的变化往往影响其他网络元素，包括路由器、安全系统、DHCP 服务器和 DNS。如果不同的个人或团队管理这些不同的元素，最好定期召开一个计划或变更控制的会议，或根据需要检查并安排即将到来的寻址的变化的计划表，以保持循环的变化带

来的潜在的影响。

一种帮助确保网络库存准确的方法，是限制那些在网络拓扑结构和 IP 寻址方案方面是权威的并有敏锐的见解的人对库存的写访问。使用一个单独的密码保护的电子表格，有且仅有一个网络管理者可以修改，这是一个避免 IP 库存不慎或错误的更改的一个方法。然而，即使是中等规模的组织，这种方法也显得相当笨拙。假设整个网络的库存依赖一个单一的个体，该个体必须夜以继日地工作，而且当他或她离开组织时，恢复对库存化的访问可能是非常困难的，除非提前培养继任者。

多个同时支持管理员，是市场上绝大多数的网络管理系统的一个重要特征，并且大多数允许一定程度的范围控制，因此特定管理员只能在特定的设备或部分网络上执行某些特定的功能。请确定你所选择的系统支持管理员日志是否应该需要调查在系统上"谁做了什么"。

与有纪律的多管理员受审视的访问网络库存同样重要的，是把责任制、任意修改委托给能在库存范围外生成的 IP 地址分配、DNS 资源记录和设备配置。例如，手动配置可能被打错，子网可能被配置在错误的路由器接口上，还有客户端或 DHCP 对 DNS 的更新都可能导致网络库存偏离现实。库存是一个网络和地址规划的模型，并且网络管理任务依赖规划的准确性。因此，额外的"脉冲读数"则需要从网络本身的获取。定期轮询和比较来自带库存的网络的实际配置，是确保库存准确的关键。

#### 9.4.1.2 网络发现

有多种方法可用来收集实际的网络数据，从 ping 到 DNS 查找，到 SNMP 轮询。执行 ping 能够检测 IP 地址的占有者，并提供了一个基本的方法，以确定与 IP 库存相应的部分进行比较哪些 IP 地址正被使用。虽然正如本书第 6 章所讨论的，这本身不是一个现实的发现方案。ping6 命令对验证一个给定目标的 IPv6 连接和寻址非常有用，但要注意，有些路由器或防火墙可能会丢弃 ping 命令数据报，或者甚至有些设备可能被配置为忽略 ping 命令操作。设置远程 ping 命令的代理来执行本地执行 ping 命令可以帮助避免路由器/防火墙穿越问题。

Nmap 能在 insecure. org/nmap 上自由访问，是一个非常有用的 IP 发现工具。它结合了几个发现机制，以便从连接到 IP 网络的设备收集各种信息，包括链路本地多播 ping 命令、直接主机 ping 命令、DNS 查找和端口扫描。在遍历子网时，Nmap 可以在一个命令中执行这些任务，包括发出 ping 命令到每个地址，在 DNS 中寻找一个相应的 PTR 记录和试图连接到不同的 TCP 和 UDP 端口来识别设备的操作系统。ping 命令的结果有助于确定 IP 地址的占用，DNS 查找帮助证实 DNS 的 IP 库存的主机名到 IP 地址的映射，此外端口扫描可以提供关于占用每个 IP 地址的设备类型的额外信息。

SNMP 是发现与库存相关的网络信息的另一个手段。虽然大多数终端设备，如笔记本电脑或 VoIP 电话本身不启用 SNMP，但大部分的基础设施元素，如路由器、交换机和服务器则会启用。在路由器 MIB 内，令人特别感兴趣的是接口、IP 地址和 ARP 表。如果你的基础设备支持 MIB-Ⅱ，则在不同的产品上，这些表的解释应该是一致的。只是要注意细微的差异，甚至是在来自同一供应商的不同产品之间的差异。通过这些表中的信息，能够收集到与由路由器报告的一样的接口和子网接口信息。这提供了对一般库存的有用的验证，但也可以在分配，移动或删除的子网和设备的过程中被轮询。

轮询路由器的邻居发现表，如“邻居”SNMP 表，可以提供一个对于给定的子网的明确的 MAC 地址到 IPv6 地址的映射。这种方法提供了比野蛮地执行 ping 命令更有效的手段来进行 IPv6 主机发现。一旦使用这些信息确定一列网络占有者，可以查询每台独立的主机更多的信息。因此，详细的 IPv6 主机发现一般需要两个步骤：确定子网中的一组主机，然后遍历每个主机以获得更多的细节。

#### 9.4.1.3 IP 库存调整

网络发现信息提供了一个实际的用以检查真实的子网分配、IP 地址的分配，以及相关的资源记录的手段。通过将发现的信息与 IP 库存数据库进行比较，可以确定和查清存在的差异。虽然这个比较可能需要“目测”库存的电子表格和发现的输出之间的差异，但有几个原因可以证明这样的努力是很有益。例如，数据库差异能被确定可能是基于以下几项：

- 不正确的路由器配置。不正确的子网、掩码、路由器的接口等。
- 不完整的路由器配置。计划的变更尚未实现。
- 设备的可达性问题。如果一个设备应该在一个给定的 IP 地址，且没有接收到响应。这可能是由于一个设备停运、一次短暂中断（重启）、地址重新分配或网络不可达。
- 不正确的 IP 地址分配。手动配置的地址不正确或设备从一个意想不到的地址池获取 DHCP 地址。
- 实际 IP 地址分配。在一些分散的情况下，子网上设备的安装程序可能会选择一个 IP 地址；发现可以相应地用来更新 IP 库存。
- 不完整的 IP 地址分配。分配过程的所有方面，无论是手动或是 DHCP，是不完整的。这个问题特别适用于手动分配的地址，配置分配的 IP 地址和更新 DNS 都需要人员的努力。
- 流氓设备的存在。一个未知的或未经授权的设备已获得一个 IP 地址。这提供了一个的有效的后访问控制机制，以补充和审计网络访问控制解决方案。

除了检测差异，分析发现信息也可以确认分发或配置任务的完成，以及删

除任务的完成。在移动块、子网和 IP 地址时，数据发现是必不可少的。由于移动需要新地址的分配、移动，然后是旧地址的删除，所以在删除旧 IP 库存的地址之前必须有移动完成的确认。移动完成之前，这些地址不应该被删除，所以它们不会在不知情的情况下，在被实际放弃之前被分配给其他设备或子网。

总之，网络发现对保证 IP 库存的准确性是必不可少的。这也有利于监测配置或分配进度和时间段，管理包含多个子任务的任务完成，以及检测不正确的分配和潜在的恶意设备。

### 9.4.2 地址回收

前面所讨论的网络发现和调整的另一个好处，是对设备可达性问题的检测。如果服务器已在一个给定的 IP 地址上设置并被回应，但现在不再如此，这样的事件应该激发进一步的调查。如果没有移动或淘汰设备的计划，或者没有达到子网其他设备上的网络问题，那么该设备可能正经历一个中断，可能是重新启动，或者可能已被移动或断开连接，又或者可能已经被重新分配地址。如果服务器正在提供关键的服务或应用，你应该通过网络管理系统监视其状态，该系统可以证实停运理论和触发纠正措施。如果 IP 地址在下一次尝试被发现，也许它只是简单地重新启动。如果它在未来 N 次尝试中均不响应，那也许它已经在物理上不复存在了（或至少是不通电了）。不幸的是，人们并不总是告知 IP 规划团队，设备已被删除或转移其他地方，即使在联系最紧密的组织中。一个迅速通知对方检查设备状态的电话可能确实有用，但要确认一个检验设备的设备所有者通常是困难而且耗时的。

然而，评估设备可能的命运的关键点是，它可能需要进行多个发现尝试，以确定设备是否曾经存在而现在不在，或只是遭遇了短暂的停电或断开，或是被借走了而现在已经返还。跟踪一系列发现的尝试，可能会很困难。运行日志或电子表格可以用来记录不符或“失踪”的 IP 地址，当它们不能被检测到的时候。随着时间的推移查看该日志，可能有助于确定一个 IP 地址是否被记录为在使用但实际上并不是。

在检验这样的日志时，如果给定 IP 地址直到一个月前被成功发现，当它在这么多的尝试（如 30 次）后最终到达，那么它可能会被证实可用于未来的分配或可收回的。回收的概念需要确定一些 IP 地址，这些 IP 地址在 IP 库存中表示为可以使用的，但在现实中没有使用，在最近的历史记录中也没有被使用。分析多个发现结果提供了更强大的样本集，回收决定是以其为基础的，本质上是把设备从库存中删除和把地址释放以分配到另一台设备。回收功能虽然是一个强大的 IPv4 网络管理的功能，但它在手动配置的 IPv6 地址之外的价值有限。

大多数设备很可能会使用带隐私扩展的 DHCPv6 或 SLAAC。除了为从 IP 库

存删除设备提供的强大确认，回收可能也同样被应用于子网。当移动或删除一个子网时，通常建议验证所有 IP 地址占有者是否已被删除，且不再使用子网[⊖]的 IP 地址。分析来自一个给定的子网上的所有地址的发现结果，可以保证子网能被删除。但是类似 IP 地址回收，多个样本集提供对于可回收处置的更强大的确认。只要记住，你很少会在一个子网上看到零响应，至少当它还配置在一个路由器接口上时，所以你要检查忽略了路由器、交换机或其他的设备类型的连续的发现结果。

## 9.5 性能管理

性能管理涉及监控网络和重要网络元素的性能。跟踪关键构件的硬件的统计数据，如 CPU 利用率、内存、磁盘和网络接口的输入/输出（I/O），是非常有用的。这种监测实现了对硬件管理运行在设备上的服务的能力的追踪。趋势分析在这方面及对未来的硬件采购的主动规划是有益的，使得可以实现更多的服务器之间的负载分配。

### 9.5.1 服务监控

监测正常运行时间和硬件统计对维持高质量的核心服务至关重要，但你可能还需要监控这些设备的协议功能。例如，监控 DNS 的服务可以帮助保证足够的 DNS 能力以满足名称解析的需求，并帮助识别任何异常情况。从客户端的角度测量去验证应用程序的功能，需要定期发布 DNS、NTP 查询或 DHCP 请求报，并测量收到一个适当的响应[⊖]需要的响应时间。此应用程序测试可能来源于部署在不同地点的服务探测仪，以产生这些“综合交易”，并测量和存储响应时间结果。分析来自不同探查的历史数据可以提供敏锐的视角，洞察到 DNS/DHCP/NTP 服务和网络性能。

### 9.5.2 应用性能管理

虽然服务通过如网络地址分配和名称解析等数据通路之外的功能支撑起网络，带内基础设施的监测对整个网络的运行状况和性能是否能维持在良好状态同样是至关重要的。这不仅包括对基础设施设备的健康状态和性能状态的监控，如交换机、路由器和负载平衡器，还包括在你的网络中“流动”的应用的健康

⊖ 忽略路由器的 IP 地址占用，因为它通常会在子网上识别自身。

⊖ 故障管理部分中提到，在没有响应的情况下，可能表明一个服务中断，且当它持续存在时，应该对其进行调查。

和性能状态的监控。通过跨网络访问的应用支撑起用户体验的视图，使得问题的主动检测、更为简单的诊断和所报告问题的解决得以实现。

添加 IPv6 到网络不一定影响应用的性能。不过，在网络中使用隧道或转换可能会妨碍用户体验，尤其是在高容量段。在进行生产部署之前，先搁置测试阶段的资源，尝试先描绘计划的隧道、翻译或甚至双协议栈的潜在性能影响。这将帮助确定哪种方法可能最好地满足你的网络用户的需求；也有利于建立一个知识库，用于解决在可能的解决路径上出现的潜在性能问题。

### 9.5.3 审计和报告

大多数管理系统一般提供一定程度的关于"谁做了什么"的审计和不同层次的报告。这些可能只是很容易被归类在计费管理的功能，使管理员能够跟踪和排查活动并以报告格式表达状态信息。IP 地址使用情况的审计，即谁在特定的时间点拥有给定 IP 地址，对于解决网络问题或调查潜在的非法活动，是很有价值的信息。同样，如果你正试图跟踪一个给定设备的 IP 地址占用的历史，由硬件地址进行报告也是有利的。

在没有网络管理系统的情况下，执行这样的审计可能是很困难的，除非是在最小的网络中。处理服务器和基础设施与日俱增的日志和警报的迭代转储是很有必要的。此过程使为了故障排除、变更控制和审计而进行的配置状态和性能的跟踪得以实现。

对 IP 地址规划感兴趣的通用报告包括以下内容，虽然你的系统可能提供不同的或额外的报告。

- 网络资产报告。网络元件和服务的逐项报告，以及配置的摘要信息。
- 地址分配报告。当前快照和/或历史记录中由子网或块分配的地址的摘要信息。
- 地址差异报告。IP 库存与发现的 IP 地址信息之间差异的重点。
- 服务性能报告。网络服务协议消息按类型的摘要和详细信息，以及（或者）客户端与服务器的关键指标摘要。
- 应用性能报告。应用程序的响应时间和负载的摘要和详细信息，以及应用程序服务器的关键指标摘要。
- 审计报告。管理员通过子网、设备、硬件地址、路由器及服务器的活动。

## 9.6 安全管理

本书第 6 章主要介绍了安全策略，这些策略应该考虑在你的 IPv4/IPv6 网络中实施。一旦投入使用，防火墙和其他安全系统必须进行监测，以检测出可能

的攻击。监测过滤的数据报能够识别使用未提供服务的攻击或尝试，如移动 IPv6。回应最终用户对于远程访问的投诉，可能需要某些政策的松动，虽然经过这样的变化，但仍需要密切监测以发现利用新机会的可能攻击者。如果你有定期的安全审查，与 IPv6 相关的结果和分析需要被添加到议事日程中。

## 9.7 灾难恢复/业务连续性

业务连续性实践努力在面对重大故障时维持组织的运转。重大故障或“灾难”，意味着故障的绝对规模超出了少量服务器或网络设备。必须提前记录自动和手动步骤以重新配置或重新调配资源，从而维持网络和应用程序（或至少是重要的服务和应用程序）的运行。核心网络元素和服务应当部署在冗余配置中，从而在路由器、服务器、应用程序或链路中断的情况下提供网络连续性。

网络业务的连续性，可能需要部署额外的网络元素、管理系统和数据库。多个活动数据库或主用/备用配置的部署将取决于你为每个组件所选择的供应商。供应商要实现多种方法来方便冗余，如全数据库复制和转移、需要某种程度的网络分区的多主数据库，以及使用存储区域网络、SQL 或 LDAP 复制能力的数据库复制技术的部署。需要执行灾难恢复的操作任务同样因每个供应商的不同而有所差异。

# 第10章　IPv6和因特网展望

从20世纪90年代最早商业化开始，因特网已经取得了极大的成功。因特网协议被证明是可升级的、健壮的、可扩展的，这也是它的发明者有先见之明的体现。而且它的成功之处还不止于此。正如本书第1章讨论的那样，在过去的十年中，这方面世界年复合增长率（Compound Average Growth Rate，CAGR）的平均值是18%。而且发展中国家在无线通信基础设施上的投资特别能促进互联网需求。

因特网这样的成功带来的结果是耗尽用于因特网通信的IP地址。如果还想因特网这样持续增长，唯一的办法是采用IPv6。还好，IPv6技术已经存在十多年了，并且已经相对成熟，而且，很多不同的供应商和通信设备都支持这个协议了。在有了IPv6的部署以后，现在的因特网是一个混合IPv4-IPv6的网络，而且IPv6用户和IPv6流量将会持续增长。IPv6是唯一的可以满足目前持续增长的因特网需求的，而那些希望和这些持续增长的IPv6用户们通信的组织必须要部署IPv6。

## 10.1　促成技术的因素

能为因特网接入新用户提供地址能力支持，是IPv6的一个重要优势。但是IPv6还有一些新的特征，而且这些特征可以促进因特网应用的新思维和因特网演化。这中间最重要的就是SLAAC和对移动性支持的提高。和IPv6一起，以下这些技术的提升为人和物与因特网的联系提供了一个好的手段：

- IPv6，特别是SLAAC和移动性。
- 计算设备的提升，特别是小型化、多媒体能力和计算能力提升及费用上的下降。
- 节约电、符合时势、安全和空气传播高效的无线通信协议。
- 可以处理大量的输入和大量传感器的管理软件为自动化的操作或人类消费而过滤和存储数据。

这些技术的改进将会不断制造出更小、更节能的智能手机、平板电脑和别的便携式用户设备。这些技术也会帮助实现因特网与“物”的连通。今天一个主要的研究课题是关于未来“物联网（Internet of Things，IoT）”的探索。IoT由各种设备和传感器组成，这些设备和传感器从人、地方和事物上收集信息并进

行聚合、处理和报告。

在任何地方、任何时间能和任何人、任何物通信的能力，给因特网用户带来的是更好的监控能力和更有效的资源利用，以使人们减少花费、保护环境甚至提供一个内心的宁静。设法利用这些技术的组织可以开发和提供更智能的服务和产品。关于智能服务和智能产品的应用的例子如下：

- 智能应用（smart application）。为更多的智能资源管理和客户服务提供一个对于一些没有察觉到的大量数据的中心视觉。这些系统包括：
  - 智能电网（smart grid）。根据需求自动调控电力、水、天然气等资源，减少资源浪费，节省消费者公用事业账单。
  - 智能汽车（smart car）。在汽车内进行诊断的和使用的传感器，可以提供性能报告，发现和恢复故障，提供已坏组件的客户通知，推荐服务检查，以及自动的碰撞探测和报告。
  - 智能家居（smart home）。对房子远程监控，远程控制用电设备，以及制热、制冷、光照、娱乐和接入设施。
  - 智慧城市（smart citie）。使用交通灯动态进行交通管理，通过留言板沟通可选路径，以及提高能量利用率。
- 市政和工业的监控和监测（municipal and industrial surveillance and monitoring）。物理访问的控制和监控、极端条件下的环境监控（如自然灾害、火灾、水灾）、结构监测和交通监测。
- 野外应用（field application）。舰船的管理、调度和交通工具的远程信息处理。
- 卫生保健（healthcare）。远程监视病人的生命特征，进行诊断和药物管理，建立“体域网（body area networks）”，严格监视医用库存环境（如存储血浆和器官的保存环境）。
- 工业（industrial）。工厂流水线监视，诊断程序，资源控制，供应链管理，操作过程的监视，以及控制无线网络提供的可达性。
- 军事（military）。战场专用网络通过不同士兵的传感器向军事指挥报告最新状态。
- 消费者（consumer）。位置感知服务（如用于跟踪小孩或宠物、市场服务（寻找最近的商店）），通过数字标识的电子支付，游戏服务，以及其他消费者服务。

## 10.2 因特网的阴暗面

跟任何使能技术一样，随着规模和实力的增长，人们的日常生活越来越依

赖因特网的接入，限制网络接入的可能性也随之增加。例如，现在很多组织都会限制一些网络流量的流入和流出。而这个“审查制度”已经被一些政府、网络服务提供商和内容提供商使用，并且还将被继续使用。控制了内容和应用分发供应链的设备制造商很可能限制到一些网站和应用的访问。宽带提供商和网络服务提供商可能会根据订购费，使流量与之符合，这和“网络中立（net neutrality）”的平等对待每个 IP 数据报的原则相矛盾[100]。

随着那些可以提供途径获取传感器、相机和其他形式信息收集设备的增长，那些被授权的组织可以加强对这些信息免费使用的控制。由于存储硬件价格的持续下降，很多组织就可以对这些“受监督”的数据进行收集、分析并使用，可能用在好的地方，也可能用在不好的地方。

在一个乌托邦式的幻想世界中，每个人可以有“平等访问（equal access）”的权利来访问因特网上发布的所有数据。但是当考虑到如相片、个人信息和银行卡数据等私人数据会被任何人访问时，马上就会明白，对访问的限制是必不可少的。但是限制的边界在不同控制等级的个人、企业法人、服务提供商和政府上，变得模糊不清了。有的人愿意相信，那些希望世界上所有人都能访问的，发布在因特网上的信息，可以真正在世界范围内被访问；而那些希望限制访问范围的信息，不会在这个范围之外被访问到。可是这两个场景都不能真正保证。

在所有这些场景中，对网络信息的限制访问，与因特网使用的具体协议版本没有关系，也不是新事物。这种对个人自由的限制，在 IPv4 和 IPv6 是一样的。

## 10.3 因特网的光明未来

不管这些阴暗面是否充斥网络社会和其他各方面，人们对未来因特网，及其在通过促进全球通信使世界变小、自动化、日常生活的远程控制和提升生活质量中所扮演的角色，持乐观态度。IPv6 明智的定位会为支持这一角色提供核心技术。

### 10.3.1 更加智能地生活

即将到来的智能家居、智能能源和智能汽车让消费者可以更好地控制能量、水资源和通信设备，为消费者提供自动化维护功能，为汽车、家庭用具和资源消耗提供更好的信息和诊断功能。这些有效的“自我提醒（self-reminding）”技术可以减少担忧并且降低花费，同时减少对环境资源的需求。

### 10.3.2 保持踪迹

在照顾对家庭、小孩、宠物或者年长亲属方面，会变得更能负担得起并且实现上也更简单。可联网的监控摄像机提供视频监控和记录，而可穿戴或其他携带传感器可以用于定位追踪。后一个应用会很受小孩父母和宠物主的欢迎。

### 10.3.3 可扩展的医疗保健

病人在完成医学治疗出院之后，通过新兴的远程病人诊断和医疗保健技术可以对刚回家病人的生命体征或其他参数进行与位置无关的监视。指标读数超过可接受的阈值时会发出通知来触发后续行动。这个技术也可以被用在不那么严格的日常监控应用中。

### 10.3.4 公共安全

对于公共安全应用，如警察、救护车、火警救援的应用，相关信息越快到达最近的相关单位越好采取行动。使用传感器信息、监控摄像机或其他触发器，信息可以被分发到最近的办公室以采取行动。这些应用也被用于灾难准备、灾难恢复和救助的行动中。

### 10.3.5 未来的信用卡

随着 IPv6 技术促进世界上智能手机的增长，这些设备的新用途将是智能手机代替信用卡这样的功能。当前一个严重的问题是，通过收集因特网信用卡交易的数据，进行身份盗用和冒用。如果信用卡被替代成一些不那么容易复制的技术产品将会怎么样？人们已经见证了从现金到借记卡/信用卡的变化，那么使用智能手机代替借记卡/信用卡就如同使用虚拟信用卡一样。

“信用卡之父”——Jermome Svigals，预测智能手机在将来的 10 年超过信用卡成为一个主要的支付工具。在他的书——《智能手机中的银行（Bank on Your Smartphone)》（即本书参考文献）[101]），预测零售业银行正在向一个新的时代演化，这个时代的中心是和因特网相连的手持智能手机。虽然是由其他协议而非 IPv6 来进行金融数据传输，但是手机数量的大量增加将会是这次演化的主要推动力，而 IPv6 将协助推动智能手机在世界上的广泛使用。

### 10.3.6 消费应用

位置感知应用（location-aware application）和轮廓感知信息（profile-aware information）可以用在零售企业中，它们可以随机地发送优惠券或者精心选择的广告，给那些经过零售店面的顾客。有的人也许会觉得这是骚扰，但是，假如

你想要喝一杯咖啡，这时候一个拿铁咖啡的打折广告出现了，并且打折商店就在一个街区之外，你一定会觉得这件事无法抗拒。

另外一个潜在的应用方向是保险业，它需要收集所有这些房屋、汽车和健康传感器的信息，用于评估用户保险费用，也用于找出事故的根本原因。

## 10.4　小结

本章所描述的应用案例仅是在增长的互联世界中使生活更美好的应用和机会中的皮毛。虽然每一个好处都会附带一些注意点，但是因为科技持续向更小、更便宜、更智能的方向发展。更重要的是，通过 IPv6 网络设备的发展，获取信息的途径将会扩大到一个前所未有的高度，而未来也将会需要有效的管理应用软件按需进行分析、报告和告警。

# 附录　IPv6 准备情况评估模板修订 1

这个模板给出了对于一个给定组织的 IPv6 准备情况评估，用于收集和提供需要获取 IPv6 准备情况的当前状态信息的中心存储库。该模板的电子版本可从 http：//www. ipamworldwide. com 下载使用。

## A. 1　IP 地址分配

本节概述了当前的 IP 地址规划。

| 功能区 | 项　　目 | 块　地　址 | 使用/利用率 | 评　　估 | 下一步的计划 |
|---|---|---|---|---|---|
| IPv4 地址空间 | | | | | |
| | ARIN 根块分配 | | | | |
| | RIPE 根块分配 | | | | |
| | APNIC 根块分配 | | | | |
| | LACNIC 根块分配 | | | | |
| | AfriNIC 根块分配 | | | | |
| | RFC 1918 根块分配 | | | | |
| | | | | | |
| IPv6 地址空间 | | | | | |
| | ARIN 根块分配 | | | | |
| | RIPE 根块分配 | | | | |
| | APNIC 根块分配 | | | | |
| | LACNIC 根块分配 | | | | |
| | AfriNIC 根块分配 | | | | |
| | ULA 根块分配 | | | | |

## A. 2　流程与人员

本节概述了对于流程、实践和人员的评估标准。

| 功能域 | 项　　目 | IPv6 就绪或认证? | 包含 IPv6 的下一步的计划 |
|---|---|---|---|
| IT 管理过程 | | | |
| | 过程 1 | | |
| | 过程 2 | | |
| | | | |
| 安全过程 | | | |
| | 过程 1 | | |
| | 过程 2 | | |
| | | | |
| 工作人员准备情况 | | | |
| | 网络架构师 1 | | |
| | 网络工程师/分析师 1 | | |
| | 网络技术员 1 | | |
| | IT 帮助台工作人员 1 | | |

| 功能区 | 项　目 | 项目 *ID* | 当前厂商 | 当前版本 | 当前硬件与操作系统 | *OS* 支持 *IPv6*? | 功能支持 *IPv6*? | 评估:选择一项 | | | *IPv6* 的具体限制 | 需要的其他单元 | 具有 *IPv6* 功能的下一步的计划 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 项目完全支持 *IPv6* | 项目支持 *IPv6* 升级 | 项目不支持 *IPv6* | | | |
| 网络服务 | | | | | | | | | | | | | |
| | DHCP | | | | | | | | | | | | |
| | DNS | | | | | | | | | | | | |
| | NTP/SNTP | | | | | | | | | | | | |
| | Radius/Oiameter | | | | | | | | | | | | |
| | FTP | | | | | | | | | | | | |
| | TFTP | | | | | | | | | | | | |
| | Rsync | | | | | | | | | | | | |
| | SMTP/POP/IMAP | | | | | | | | | | | | |
| | HTTP | | | | | | | | | | | | |
| | 其他 | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
| 网络基础设施 | | | | | | | | | | | | | |
| | 路由器 | | | | | | | | | | | | |
| | 核心交换机 | | | | | | | | | | | | |
| | 分布式/边缘交换机 | | | | | | | | | | | | |
| | 负载均衡器 | | | | | | | | | | | | |
| | 应用服务器 | | | | | | | | | | | | |
| | 防火墙 | | | | | | | | | | | | |
| | *SAN/NAS* 存储系统 | | | | | | | | | | | | |
| | 无线访问点 | | | | | | | | | | | | |

（续）

| 功能区 | 项　　目 | 项目 *ID* | 当前厂商 | 当前版本 | 当前硬件与操作系统 | *OS* 支持 *IPv6*? | 功能支持 *IPv6*? | 评估：选择一项 | | | *IPv6* 的具体限制 | 需要的其他单元 | 具有 *IPv6* 功能的下一步的计划 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 项目完全支持 *IPv6* | 项目支持 *IPv6* 升级 | 项目不支持 *IPv6* | | | |
| 网络基础设施 | | | | | | | | | | | | | |
| | *IP* 电话服务器 | | | | | | | | | | | | |
| | 其他 | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
| 终端用户/端点系统 | | | | | | | | | | | | | |
| | 台式机/笔记本/工作站 | | | | | | | | | | | | |
| | 平板电脑/*PDA* | | | | | | | | | | | | |
| | 智能手机 | | | | | | | | | | | | |
| | 其他手持设备 | | | | | | | | | | | | |
| | 打印机 | | | | | | | | | | | | |
| | 销售点设备 | | | | | | | | | | | | |
| | *CPE* 设备 | | | | | | | | | | | | |
| | 其他 | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
| 软件应用程序 | | | | | | | | | | | | | |
| | 商业应用 1 | | | | | | | | | | | | |
| | 网络管理应用 1 | | | | | | | | | | | | |
| | *OSS* 应用 1 | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
| 客户/合作伙伴/链接 | | | | | | | | | | | | | |
| | 合作伙伴系统 1 | | | | | | | | | | | | |
| | 合作伙伴系统 2 | | | | | | | | | | | | |

# 参考文献


1. Internet World Stats. *Internet Usage Statistics.* Internet World Stats. [Online] Miniwatts Marketing Group, June 30, 2012. [Cited: January 19, 2013.] http://www.internetworldstats.com/stats.htm.
2. Kim, Y., Kelly, T., Raja, S. *Building Broadband: Strategies and Policies for the Developing World.* s.l.: World Bank, January 2010.
3. Netcraft. 2012*archives.* Netcraft. [Online] http://news.netcraft.com/archives/2012/.
4. Hubbard, K., Kosters, M., Conrad, D., Karrenberg, D., Postel, J. *Internet Registry IP Allocation Guidelines.* s.l.: IETF, November 1996. RFC 2050.
5. International Monetary Fund.*World Economic Outlook (WEO).* October, 2012.
6. *Eastern Europe continues to move up the broadband and IPTV league tables,* Press release, London, UK: Point Topic Ltd., October, 2012.
7. The World Bank. *GDP Growth.* The World Bank. [Online] [Cited: January 21, 2013.] http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG.
8. RIPE NCC. *IPv6 Enabled Networks.* RIPE NCC. [Online] http://v6asns.ripe.net/v/6?s=_ALL;s=_RIR_APNIC;s=_RIR_AfriNIC;s=_RIR_ARIN;s=_RIR_LACNIC;s=_RIR_RIPE_NCC.
9. Kim, E., Kaspar, D., Vasseur, JP. *Design and Application Spaces for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs).* s.l.: IETF, April 2012. RFC 6568.
10. Delgrossi, L., Berger, L., ed., Internet Stream Protocol Version 2 (ST2) Protocol Specification - Version ST2+, s.l.: IETF, August, 1995, RFC 1819.
11. Rooney, T. *IP Address Management Principles and Practice.* Hoboken, NJ: Wiley, 2011.
12. Deering, S., Hinden, R. *Internet Protocol, Version 6 (IPv6) Specification.* s.l.: IETF, December 1998. RFC 2460.
13. Protocol Numbers. *IANA.* [Online] http://www.iana.org/assignments/protocol-numbers/protocol-numbers.xml.
14. Kawamura, S., Kawashima, M. *A Recommendation for IPv6 Address Text Representation.* s.l.: IETF, August 2010. RFC 5952.
15. Hinden, R., Deering, S. *IP Version 6 Addressing Architecture.* s.l.: IETF, February 2006. RFC 4291.
16. Kohno, M., Nitzan, B., Bush, R., Matsuzaki, Y., Colitti, L., Narten, T. *Using 127-Bit IPv6 Prefixes on Inter-Router Links.* s.l.: IETF, April 2011. RFC 6164.
17. Internet Assigned Numbers Authority (IANA). *Internet Protocol Version 6 Address Space.* www.iana.org. [Online] [Cited: October 12, 2012.] http://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xhtml.
18. Hinden, R., Deering, S., Nordmark, E. *IPv6 Global Unicast Address Format.* s.l.: IETF, August 2003. RFC 3587.
19. Hinden, R., Haberman, B. *Unique Local IPv6 Unicast Addresses.* s.l.: IETF, October 2005. RFC 4193.
20. Haberman, B., Thaler, D. *Unicast-Prefix-based IPv6 Multicast Addresses.* s.l.: IETF, August 2002. RFC 3306.

21. Park, J.-S., Shin, M.-K., Kim, H.-J. *A Method for Generating Link-Scoped IPv6 Multicast Addresses.* s.l.: IETF, April 2006. RFC 4489.
22. Conta, A., Deering, S., Gupta, M., Eds. *Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification.* s.l.: IETF, March 2006. RFC 4443.
23. IANA. *Internet Control Message Protocol version 6 (ICMPv6) Parameters.* IANA. [Online] http://www.iana.org/assignments/icmpv6-parameters.
24. Deering, S., Fenner, W., Haberman, B. *Multicast Listener Discovery (MLD) for IPv6.* s.l.: IETF, October 1999. RFC 2710.
25. Vida, R., Costa, L. *Multicast Listener Discovery Version 2 (MLDv2) for IPv6.* s.l.: IETF, June 2004. RFC 3810.
26. Crawford, M. *Router Renumbering for IPv6.* s.l.: IETF, August 2000. RFC 2894.
27. Crawford, M., Haberman, B., Eds. *IPv6 Node Information Queries.* s.l.: IETF, August 2006. RFC 4620.
28. Rooney, T. *Introduction to IP Address Management.* IEEE Press/Wiley, 2010.
29. Narten, T., Draves, R., Krishnan, S. *Privacy Extensions for Stateless Address Autoconfiguration in IPv6.* s.l.: IETF, September 2007. RFC 4941.
30. Microsoft. *IPv6 Address Autoconfiguration.* www.microsoft.com. [Online] [Cited: October 19, 2009.] http://msdn.microsoft.com/en-us/library/aa917171.aspx.
31. Johnson, D., Deering, S. *Reserved IPv6 Subnet Anycast Addresses.* s.l.: IETF, March 1999. RFC 2526.
32. Loughney, J., Ed. *IPv6 Node Requirements.* s.l.: IETF, April 2006. RFC 4294.
33. Chown, T. *Use of VLANs for IPv4-IPv6 Coexistence in Enterprise Networks.* s.l.: IETF, June 2006. RFC 4554.
34. Thaler, D., Draves, R., Matsumoto, A., Chown, T. *Default Address Selection for Internet Protocol Version 6 (IPv6).* s.l.: IETF, September 2012. RFC 6724.
35. Wing, D., Yourtchenko, A. *Happy Eyeballs: Success with Dual-Stack Hosts.* s.l.: IETF, April 2012. RFC 6555.
36. Durand, A., Ihren, J. *DNS IPv6 Transport Operational Guidelines.* s.l.: IETF, September 2004. RFC 3901.
37. Chown, T., Venaas, S., Strauf, C. *Dynamic Host Configuration Protocol (DHCP): IPv4 and IPv6 Dual-Stack Issues.* s.l.: IETF, May 2006. RFC 4477.
38. Huston, G. *The ISP Column.* www.potaroo.net. [Online] May 2012. http://www.potaroo.net/ispcol/2012-05/notquite.html.
39. Rooney, T. *IPv4-to-IPv6 Transition and Co-Existence Strategies.* Santa Clara, CA: BT INS, Inc., March 2008.
40. Blanchet, M., Parent, F. *IPv6 Tunnel Broker with Tunnel Setup Protocol (TSP).* s.l.: IETF, February 2010. RFC 5572.
41. Huitema, C. *Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs).* s.l.: IETF, February 2006. RFC 4380.
42. Thaler, D., *Teredo Extensions.* s.l.: IETF, January, 2011, RFC 6081.
43. Ibid.
44. Thaler, D., Krishnan, S., Hoagland, J. *Teredo Security Updates.* s.l.: IETF, September 2010. RFC 5991.
45. Bound, J., Toutain, L., Richier, J.L. *Dual Stack IPv6 Dominant Transition Mechanism (DSTM).* s.l.: IETF, October 2005. draft-bound-dstm-exp-04.txt.

46. Baker, F., Li, X., Bao, C., Yin, K. *Framework for IPv4/IPv6 Translation*. s.l.: IETF, April 2011. RFC 6144.
47. Li, X., Bao, C., Baker, F. *IP/ICMP Translation Algorithm*. s.l.: IETF, April 2011. RFC 6145.
48. Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., Li, X. *IPv6 Addressing of IPv4/IPv6 Translators*. s.l.: IETF, October 2010. RFC 6052.
49. Huang, B., Deng, H., Savolainen, T. *Dual-Stack Hosts Using "Bump-in-the-Host" (BIH)*. s.l.: IETF, February 2012. RFC 6535.
50. Tsuchiya, K., Higuchi, H., Atarashi, Y. *Dual Stack Hosts using the "Bump-in-the-Stack" Technique (BIS)*. s.l.: IETF, February 2000. RFC 2767.
51. Lee, S., Shin, M.-K., Kim, Y.-J., Nordmark, E., Durand, A. *Dual Stack Hosts Using "Bump-in-the-API" (BIA)*. s.l.: IETF, October 2002. RFC 3338.
52. Bagnulo, M., Matthews, P., van Beijum, I. *Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers*. s.l.: IETF, April 2011. RFC 6146.
53. Tsirtsis, G., Srisuresh, P. *Network Address Translation - Protocol Translation (NAT-PT)*. s.l.: IETF, February 2000. RFC 2766.
54. Aoun, C. Davies, E. *Reasons to Move the Network Address Translator—Protocol Translator (NAT-PT) to Historic Status*. s.l.: IETF, July 2007. RFC 4966.
55. Leech, M., Ganis, M., Lee, Y., Kuris, R., Koblas, D., Jones, L. *SOCKS Protocol Version 5*. s.l.: IETF, March 1996. RFC 1928.
56. Kitamura, H. *A SOCKS-based IPv6/IPv4 Gateway Mechanism*. s.l.: IETF, April 2001. RFC 3089.
57. Hagino, J., Yamamoto, K. *An IPv6-to-IPv4 Transport Relay Translator*. s.l.: IETF, June 2001. RFC 3142.
58. Vincent, M. *Vin's World*. Vin's World. [Online] http://vinsworldcom.blogspot.com/2011/12/cat-with-10-lives.html.
59. Rooney, T. *Service Provider IPv6 Deployment Strategies*. s.l.: BT INS Inc., 2011.
60. Yamagata, I., Shirasaki, Y., Nakagawa, A., Yamaguchi, J., Ashida, H. *NAT444*. s.l.: IETF, January 2011. draft-shirasaki-nat444-03.txt.
61. Donley, C., Howard, L., Kuarsingh, V., Chandrasekaran, A., Ganti, V. *Assessing the Impact of NAT444 on Network Applications*. s.l.: IETF, October 2010. draft-donley-nat444-impacts-01.txt.
62. De Clercq, J., Ooms, D., Prevost, S., Le Faucheur, F. *Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)*. s.l.: IETF, February 2007. RFC 4798.
63. De Clercq, J., Ooms, D., Carugi, M., Le Faucheur, F. *BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN*. s.l.: IETF, September 2006. RFC 4659.
64. Despres, R. *IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)*. s.l.: IETF, January 2010. RFC 5569.
65. Townsley, W., Troan, O. *IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)—Protocol Specification*. s.l.: IETF, August 2010. RFC 5969.
66. Rooney, T. *IP Address Management Principles and Practice*. Hoboken: IEEE Press, 2011.
67. Wu., J., Cui, Y., Li, X., Xu, M., Metz, C. *4over6 Transit Solution Using IP Encapsulation and MP-BGP Extensions*. s.l.: IETF, March 2010. RFC 5747.
68. Durand, A., Droms, R., Woodyatt, J., Lee, Y. *Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion*. s.l.: IETF, August 2011, RFC 6333.

69. Internet Society. *Internet Society*. [Online] http://www.internetsociety.org/news/internet-society-number-resource-organization-and-regional-internet-registries-reinforce.
70. Netformx Discovery. *Netformx*. [Online] http://www.netformx.com/discovery.
71. HP DDMI. *HP*. [Online] http://www8.hp.com/lamerica_nsc_carib/en/software/software-product.html?compURI=tcm:246-936991.
72. OPNET NetMapper. *OPNET*. [Online] http://www.opnet.com/solutions/network_management/netmapper.html.
73. IANA. *Number Resources.* www.iana.org. [Online] [Cited: October 20, 2009.] http://www.iana.org/numbers/.
74. AfriNIC. AfriNIC Home Page. *www.afrinic.net*. [Online] [Cited: October 20, 2009.] http://www.afrinic.net/.
75. APNIC. APNIC Home Page. *www.apnic.net*. [Online] [Cited: October 20, 2009.] http://www.apnic.net/.
76. ARIN. ARIN Home Page. *www.arin.net*. [Online] [Cited: October 20, 2009.] http://www.arin.net/.
77. LACNIC. LACNIC Home Page. *www.lacnic.net*. [Online] http://www.lacnic.net/.
78. RIPE NCC. *RIPE Network Coordination Centre Home Page.* www.ripe.net. [Online] [Cited: October 20, 2009.] http://www.ripe.net/.
79. Huitema, C. *The H Ratio for Address Assignment Efficiency*. s.l.: IETF, November 1994. RFC 1715.
80. Durand, A., Huitema, C. *The Host-Density Ratio for Address Assignment Efficiency: An Update on the H Ratio*. s.l.: IETF, November 2001. RFC 3194.
81. Rooney, T. *IPv6 Addressing and Management Challenges*. Santa Clara, CA: BT INS, Inc., March, 2008.
82. Blanchet, M. *A Flexible Method for Managing the Assignment of Bits of an IPv6 Address Block*. s.l.: IETF, April 2003. RFC 3531.
83. Wasserman, M., Baker, F. *IPv6-to-IPv6 Network Prefix Translation*. June, 2011. RFC 6296.
84. Bates, T., Rekhter, Y. *Scalable Support for Multi-home Multi-provider Connectivity*. s.l.: IETF, January 1998. RFC 2260.
85. Abley, J., Lindqvist, K., Davies, E., Black, B., Gill, V. *IPv4 Multihoming Practices and Limitations*. s.l.: IETF, July 2005. RFC 4116.
86. Huston, G. *Architectural Approaches to Multi-homing for IPv6*. s.l.: IETF, September 2005. RFC 4177.
87. Nordmark, E., Bagnulo, M. *Shim6: Level 3 Multihoming Shim Protocol for IPv6*. s.l.: IETF, June 2009, RFC 5533.
88. ISO/IEC. *Information Technology—Open Systems Interconnection—Basic Reference Model: The Basic Model*, second edition. Genera, Switzerland: ISO/IEC, November 1994. ISO/IEC 7498-1:1994(E).
89. Internet Protocol Version 6 Address Space. *Internet Assigned Numbers Authority*. [Online] October 29, 2010. [Cited: April 12, 2012.] http://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xml.
90. IPv6 Global Unicast Address Assignments. *Internet Assigned Numbers Authority*. [Online] August 27, 2008. [Cited: April 12, 2012.] http://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.xml.
91. Davies, E., Mohacsi, J. *Recommendations for Filtering ICMPv6 Messages in Firewalls*. s.l.: IETF, May 2007. RFC 4890.

92. Krishnan, S., Woodyatt, J., Kline, E., Hoagland, J., Bhatia, M. *A Uniform Format for IPv6 Extension Headers*. s.l.: IETF, April 2012. RFC 6564.
93. Chown, T. *IPv6 Implications for Network Scanning*. s.l.: IETF, March 2008. RFC 5157.
94. Levy-Abegnoli, E., Van de Velde, G., Popoviciu, C., Mohacsi, J. *IPv6 Router Advertisement Guard*. February 2011. RFC 6105.
95. Davies, E., Krishnan, S., Savola, P. *IPv6 Transition/Coexistence Security Considerations*. s.l.: IETF, September 2007. RFC 4942.
96. Bhatia, M., Manral, V., Lindem, A. *Supporting Authentication Trailer for OSPFv3*. February 2012. RFC 6506.
97. Savola, P., Patel, C. *Security Considerations for 6to4*. s.l.: IETF, December 2004. RFC 3964.
98. Daniele, M., Haberman, B., Routhier, S., Schoenwaelder, J. *Textual Conventions for Internet Network Addresses*. s.l.: IETF, February 2005. RFC 4001.
99. Baker, F., Lear, E., Droms, R. *Procedures for Renumbering an IPv6 Network without a Flag Day*. s.l.: IETF, September 2005. RFC 4192.
100. The Economist. *The Future of the Internet—A Virtual Counter-Revolution*. The Economist. [Online] http://www.economist.com/node/16941635.
101. Svigals, J. *Bank on Your Smartphone*. Xlibris, 2012.

# IEEE 出版社系列之网络管理图书

这一系列图书的目标是为世界各地的通信与信息技术专业协会、私营机构和政府组织及研究中心，提供网络和服务管理方面的高质量的技术参考书和教科书。该系列图书注重于故障、配置、计费、性能和安全的管理，包括但不限于电信网络与服务、技术和实现、IP 网络与服务及无线网络与服务领域。

图书编辑
Thomas Plevyak
Veli Sahin

1. *Telecommunications Network Management into the 21st Century*
   Edited by Thomas Plevyak and Salah Aidarous
2. *Telecommunications Network Management: Technologies and Implementations*
   Edited by Thomas Plevyak and Salah Aidarous
3. *Fundamentals of Telecommunications Network Management*
   Lakshmi Raman
4. *Security for Telecommunications Management Network*
   Moshe Rozenblit
5. *Integrated Telecommunications Management Solutions*
   Graham Chen and Quinzheng Kong
6. *Managing IP Networks: Challenges and Opportunities*
   Thomas Plevyak and Salah Aidarous
7. *Next-Generation Telecommunications Networks, Services, and Management*
   Edited by Thomas Plevyak and Veli Sahin
8. *Introduction to IT Address Management*
   Timothy Rooney
9. *IP Address Management: Principles and Practices*
   Timothy Rooney
10. *Telecommunications System Reliability Engineering, Theory, and Practice*
    Mark L. Ayers
11. *IPv6 Deployment and Management*
    Michael Dooley and Timothy Rooney

# 国际视野　科技前沿

## 国际信息工程先进技术译丛

《IPv6部署和管理》
《虚拟网络——下一代互联网的多元化方法》
《下一代融合网络理论与实践》
《认知视角下的无线传感器网络》
《移动云计算：无线、移动及社交网络中分布式资源的开发利用》
《Android系统安全与攻防》
《内容分发网络》
《计算机网络仿真OPNET实用指南》
《移动无线信道》（原书第2版）
《LTE-Advanced：面向IMT-Advanced的3GPP解决方案》
《声学成像技术及工程应用》
《认知无线电通信与组网：原理与应用》
《LTE/SAE网络部署实用指南》
《网络性能分析原理与应用》
《云连接与嵌入式传感系统》
《IP地址管理原理与实践》
《自组织网络：GSM，UMTS和LTE的自规划、自优化和自愈合》
《实现吉比特传输的60GHz无线通信技术》
《LTE自组织网络（SON）：高效的网络管理自动化》
《UMTS中的LTE：向LTE-Advanced演进》（原书第2版）
《无线传感器及执行器网络》
《UMTS中的WCDMA－HSPA演进及LTE》（原书第5版）
《认知无线电网络》
《网络融合——服务、应用、传输和运营支撑》
《UMTS中的LTE:基于OFDMA和SC-FDMA的无线接入》
《高性能微处理器电路设计》
《吉规模集成电路互连工艺及设计》
《高级电子封装》（原书第2版）
《基于4G系统的移动服务技术》
《移动无线传感器网——技术、应用和发展方向》
《UMTS蜂窝系统的Q0S与Q0E管理》
《UMTS-HSDPA系统的TCP性能》
《基于射频工程的UMTS空中接口设计与网络运行》
《未来UMTS的体系结构与业务平台：全IP的3GCDMA网络》
《环境网络：支持下一代无线业务的多域协同网络》
《基于蜂窝系统的IMS—融合电信领域的V0IP演进》
《蜂窝网络高级规划与优化 2G/2.5G/3G/——向4G的演进》
《微电子技术原理、设计与应用》
《多电压CMOS电路设计》
《P2P系统及其应用》
《IPTV与网络视频：拓展广播电视的应用范围》
《下一代无线系统与网络》



机械工业出版社微信服务号




WILEY



上架指导　工业技术/通信工程/网络协议




ISBN 978-7-111-48725-8　定价：58.00元